

БИЗНЕС- ИНФОРМАТИКА

НАУЧНЫЙ ЖУРНАЛ НИУ ВШЭ

СОДЕРЖАНИЕ

В.Л. Макаров, А.Р. Бахтизин, Г.Л. Бекларян, А.С. Акопов
Цифровой завод: методы дискретно-событийного моделирования и оптимизации производственных характеристик7

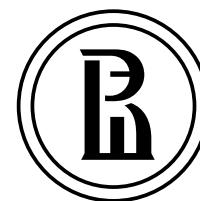
Т.К. Богданова, Л.В. Жукова
Оценка состояния объекта управления на основе универсального комплексного индикатора с использованием структурированных и неструктурированных данных21

О.И. Долганова
Улучшение клиентского опыта взаимодействия с искусственным интеллектом путем соблюдения этических принципов34

В.А. Комаров, А.В. Сарафанов
Системы интернета вещей в процессе мультидисциплинарной подготовки кадров для цифровой экономики и их проектирование47

Ф.В. Краснов, И.С. Смазневич, Е.Н. Баскакова
Проблема потери решений в задаче поиска схожих документов: Применение терминологии при построении векторной модели корпуса60

N.G. Azyabi
How do information technology and knowledge management affect SMEs' responsiveness to the coronavirus crisis?75



Издатель:
Национальный
исследовательский университет
«Высшая школа экономики»

Подписной индекс
Объединенного каталога
«Пресса России» – Е79128
Выпускается ежеквартально

Журнал включен в Перечень
русских рецензируемых
научных журналов,
в которых должны быть
опубликованы основные научные
результаты диссертаций
на соискание ученых степеней
доктора и кандидата наук

Главный редактор
Е.А. Кучерявый

Заместитель главного редактора
Е.П. Зараменских

Компьютерная верстка
О.А. Богданович

Администратор веб-сайта
И.И. Хрусталева

Адрес редакции:
119049, г. Москва,
ул. Шаболовка, д. 26-28
Тел./факс: +7 (495) 772-9590 *28509
<http://bijournal.hse.ru>
E-mail: bijournal@hse.ru

За точность приведенных сведений
и содержание данных,
не подлежащих открытой публикации,
несут ответственность авторы

При перепечатке ссылка на журнал
«Бизнес-информатика» обязательна

Тираж:
русскоязычная версия – 150 экз.,
англоязычная версия – 150 экз.,
онлайн-версия на русском и английском –
свободный доступ

Отпечатано в типографии НИУ ВШЭ
г. Москва, Измайловское шоссе, д. 44, стр. 2

© Национальный
исследовательский университет
«Высшая школа экономики»

«**Б**изнес-информатика» — рецензируемый междисциплинарный научный журнал, выпускаемый с 2007 года Национальным исследовательским университетом «Высшая школа экономики» (НИУ ВШЭ). Администрирование журнала осуществляется Высшей школой бизнеса НИУ ВШЭ. Журнал выпускается ежеквартально.

Миссия журнала — развитие бизнес-информатики как новой области информационных технологий и менеджмента. Журнал осуществляет распространение последних разработок технологического и методологического характера, способствует развитию соответствующих компетенций, а также обеспечивает возможности для дискуссий в области применения современных информационно-технологических решений в бизнесе, менеджменте и экономике.

Журнал публикует статьи по следующей тематике: моделирование социальных и экономических систем, цифровая трансформация бизнеса, управление инновациями, информационные системы и цифровые технологии в бизнесе, анализ данных и системы бизнес-интеллекта, математические методы и алгоритмы бизнес-информатики, моделирование и анализ бизнес-процессов, поддержка принятия управленческих решений.

В соответствии с решением президиума Высшей аттестационной комиссии Российской Федерации журнал включен в Перечень российских рецензируемых научных журналов, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученых степеней доктора

и кандидата наук. Журнал «Бизнес-информатика» включен в Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученых степеней кандидата и доктора наук (Перечень ВАК) по следующим научным специальностям и соответствующим отраслям науки: 05.13.10 — Управление в социальных и экономических системах (технические науки); 05.13.11 — Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей (технические науки); 05.13.18 — Математическое моделирование, численные методы и комплексы программ (технические науки); 05.25.05 — Информационные системы и процессы (технические науки); 08.00.05 — Экономика и управление народным хозяйством (по отраслям и сферам деятельности) (экономические науки); 08.00.13 — Математические и инструментальные методы экономики (экономические науки).

Журнал входит в базы Web of Science Emerging Sources Citation Index (WoS ESCI), Russian Science Citation Index на платформе Web of Science (RSCI), EBSCO.

Журнал зарегистрирован Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций (Роскомнадзор), свидетельство ПИ № ФС77-66609 от 08 августа 2016 г.

Международный стандартный серийный номер (ISSN): 1998-0663 (на русском), 2587-814X (на английском).

Главный редактор: Кучерявый Евгений Андреевич.

РЕДАКЦИОННАЯ КОЛЛЕГИЯ

ГЛАВНЫЙ РЕДАКТОР

Кучерявый Евгений Андреевич
Университет Тампере, Тампере, Финляндия

ЗАМЕСТИТЕЛЬ ГЛАВНОГО РЕДАКТОРА

Зараменских Евгений Петрович
Национальный исследовательский университет
«Высшая школа экономики», Москва, Россия

ЧЛЕНЫ РЕДКОЛЛЕГИИ

Абдульраб Абиб
Национальный институт прикладных наук, Руан, Франция

Авдошин Сергей Михайлович
Национальный исследовательский университет
«Высшая школа экономики», Москва, Россия

Акопов Андраник Сумбатович
Национальный исследовательский университет
«Высшая школа экономики», Москва, Россия

Алескерев Фуад Тагиевич
Национальный исследовательский университет
«Высшая школа экономики», Москва, Россия

Афанасьев Александр Петрович
Институт проблем передачи информации им. А.А. Харкевича РАН,
Москва, Россия

Афанасьев Антон Александрович
Центральный экономико-математический институт РАН,
Москва, Россия

Бабкин Эдуард Александрович
Национальный исследовательский университет
«Высшая школа экономики», Нижний Новгород, Россия

Баландин Сергей Игоревич
Ассоциация FRUCT, Хельсинки, Финляндия

Баранов Александр Павлович
Главный научно-исследовательский вычислительный центр
Федеральной налоговой службы, Москва, Россия

Барахнин Владимир Борисович
Федеральный исследовательский центр информационных
и вычислительных технологий, Новосибирск, Россия

Беккер Йорг
Университет Мюнстера, Мюнстер, Германия

Белов Владимир Викторович
Рязанский государственный радиотехнический университет,
Рязань, Россия

Вестнер Маркус
Регенсбургский университет прикладных наук,
Регенсбург, Германия

Гаврилова Татьяна Альбертовна
Санкт-Петербургский государственный университет,
Санкт-Петербург, Россия

Глотен Эрве
Тулонский университет, Ла-Гард, Франция

Голосов Алексей Олегович
Компания «ФОРС — Центр разработки», Москва, Россия

Грибов Андрей Юрьевич
Компания «КиберПлат», Москва, Россия

Громов Александр Игоревич
Национальный исследовательский университет
«Высшая школа экономики», Москва, Россия

Гурвич Владимир Александрович
Раттерский университет (Университет Нью-Джерси),
Раттерс, США

Демидова Лилия Анатольевна
Рязанский государственный радиотехнический университет,
Рязань, Россия

Джейкобс Лоренц
Университет Цюриха, Цюрих, Швейцария

Дискин Иосиф Евгеньевич
Всероссийский центр изучения общественного мнения,
Москва, Россия

Ефимушкин Владимир Александрович
Центральный научно-исследовательский институт связи,
Москва, Россия

Зандкуль Курт
Университет Росток, Росток, Германия

Иванников Александр Дмитриевич
Институт проблем проектирования в микроэлектронике РАН,
Москва, Россия

Ильин Николай Иванович
Федеральная служба охраны Российской Федерации, Москва, Россия

Исаев Дмитрий Валентинович
Национальный исследовательский университет
«Высшая школа экономики», Москва, Россия

Калягин Валерий Александрович
Национальный исследовательский университет
«Высшая школа экономики», Нижний Новгород, Россия

Кравченко Татьяна Константиновна
Национальный исследовательский университет
«Высшая школа экономики», Москва, Россия

Кузнецов Сергей Олегович
Национальный исследовательский университет
«Высшая школа экономики», Москва, Россия

Лугачев Михаил Иванович
Московский государственный университет
им. М.В. Ломоносова, Москва, Россия

Лин Квей-Жей
Технологический институт Нагои, Нагоя, Япония

Мальцева Светлана Валентиновна
Национальный исследовательский университет
«Высшая школа экономики», Москва, Россия

Мейор Питер
Комиссия ООН по науке и технологиям, Женева, Швейцария

Миркин Борис Григорьевич
Национальный исследовательский университет
«Высшая школа экономики», Москва, Россия

Моттль Вадим Вячеславович
Тульский государственный университет, Тула, Россия

Назаров Дмитрий Михайлович
Уральский государственный экономический университет,
Екатеринбург, Россия

Пальчунов Дмитрий Евгеньевич
Новосибирский государственный университет, Новосибирск, Россия

Пардалос Панайот (Панос)
Университет Флориды, Гейнсвилл, США

Пастор Оскар
Политехнический университет Валенсии, Валенсия, Испания

Посегга Йохим
Университет Пассау, Пассау, Германия

Самуйлов Константин Евгеньевич
Российский университет дружбы народов, Москва, Россия

Сюняев Али Рашидович
Технологический институт Карлсруэ, Карлсруэ, Германия

Таратухин Виктор Владимирович
Университет Мюнстера, Мюнстер, Германия

Триболе Жозе
Университет Лиссабона, Лиссабон, Португалия

Ульянов Михаил Васильевич
Институт проблем управления им В.А. Трапезникова РАН,
Москва, Россия

Ускенбаева Раиса Кабиевна
Международный университет информационных технологий,
Алматы, Казахстан

Цуканова Ольга Анатольевна
Санкт-Петербургский национальный исследовательский
университет информационных технологий, механики и оптики,
Санкт-Петербург, Россия

Чхартишвили Александр Гедеванович
Институт проблем управления им В.А. Трапезникова РАН,
Москва, Россия

Шмидт Юрий Давыдович
Дальневосточный федеральный университет, Владивосток, Россия

Штраус Кристина
Университет Вены, Вена, Австрия

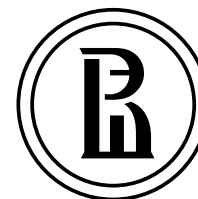
ISSN 1998-0663 (print), ISSN 2587-8166 (online)

English version: ISSN 2587-814X (print), ISSN 2587-8158 (online)

Vol. 15 No 2 – 2021

BUSINESS INFORMATICS

HSE SCIENTIFIC JOURNAL



Publisher:

National Research University
Higher School of Economics

The journal is published quarterly

The journal is included
into the list of peer reviewed
scientific editions established
by the Supreme Certification
Commission of the Russian Federation

Editor-in-Chief:

Y. Koucheryav

Deputy Editor-in-Chief

E. Zaramenskikh

Computer Making-up:

O. Bogdanovich

Website Administration:

I. Khrustaleva

Address:

26-28, build. 4, Shablovka Street
Moscow 119049, Russia

Tel./fax: +7 (495) 772-9590 *26311

<http://bijournal.hse.ru>

E-mail: bijournal@hse.ru

Circulation:

English version – 150 copies,
Russian version – 150 copies,
online versions in English and Russian –
open access

Printed in HSE Printing House
44, build. 2, Izmaylovskoye Shosse,
Moscow, Russia

© National Research University
Higher School of Economics

CONTENTS

*V.L. Makarov, A.R. Bakhtizin, G.L. Beklaryan,
A.S. Akopov*

Digital plant: methods of discrete-event modeling
and optimization of production characteristics7

T.K. Bogdanova, L.V. Zhukova

Valuating the position of the control object
based on a universal complex indicator using
structured and unstructured data21

O.I. Dolganova

Improving customer experience with artificial
intelligence by adhering to ethical principles34

V.A. Komarov, A.V. Sarafanov

IoT systems in the process of multidisciplinary
training of personnel for the digital economy
and their design47

F.V. Krasnov, I.S. Smaznevich, E.N. Baskakova

The problem of loss of solutions in the task
of searching similar documents: Applying terminology
in the construction of a corpus vector model.....60

N.G. Azyabi

How do information technology and knowledge
management affect SMEs' responsiveness
to the coronavirus crisis?75

ABOUT THE JOURNAL

Business Informatics is a peer reviewed interdisciplinary academic journal published since 2007 by National Research University Higher School of Economics (HSE), Moscow, Russian Federation. The journal is administered by HSE Graduate School of Business. The journal is published quarterly.

The mission of the journal is to develop business informatics as a new field within both information technologies and management. It provides dissemination of latest technical and methodological developments, promotes new competences and provides a framework for discussion in the field of application of modern IT solutions in business, management and economics.

The journal publishes papers in the areas of, but not limited to: modeling of social and economic systems, digital transformation of business, innovation management, information systems and technologies in business, data analysis and business intelligence systems, mathematical methods and algorithms of business informatics, business processes modeling and analysis, decision support in management.

The journal is included into the list of peer reviewed scientific editions established by the Supreme Certification Commission of the Russian Federation.

The journal is included into Web of Science Emerging Sources Citation Index (WoS ESCI), Russian Science Citation Index on the Web of Science platform (RSCI), EBSCO.

International Standard Serial Number (ISSN): 2587-814X (in English), 1998-0663 (in Russian).

Editor-in-Chief: Dr. Yevgeni A. Koucheryavy.

EDITORIAL BOARD

EDITOR-IN-CHIEF

Yevgeni A. Koucheryav
Tampere University, Finland

DEPUTY EDITOR-IN-CHIEF

Evgeny P. Zaramenskikh
National Research University Higher School of Economics,
Moscow, Russia

EDITORIAL BOARD

Habib Abdulrab
National Institute of Applied Sciences, Rouen, France

Sergey M. Avdoshin
National Research University Higher School of Economics,
Moscow, Russia

Andranik S. Akopov
National Research University Higher School of Economics,
Moscow, Russia

Fuad T. Aleskerov
National Research University Higher School of Economics,
Moscow, Russia

Alexander P. Afanasyev
Institute for Information Transmission Problems (Kharkevich
Institute), Russian Academy of Sciences, Moscow, Russia

Anton A. Afanasyev
Central Economics and Mathematics Institute, Russian Academy
of Sciences, Moscow, Russia

Eduard A. Babkin
National Research University Higher School of Economics,
Nizhny Novgorod, Russia

Sergey I. Balandin
Finnish-Russian University Cooperation in Telecommunications
(FRUCT), Helsinki, Finland

Vladimir B. Barakhnin
Federal Research Center of Information and Computational
Technologies, Novosibirsk, Russia

Alexander P. Baranov
Federal Tax Service, Moscow, Russia

Jörg Becker
University of Munster, Munster, Germany

Vladimir V. Belov
Ryazan State Radio Engineering University, Ryazan, Russia

Alexander G. Chkhartishvili
V.A. Trapeznikov Institute of Control Sciences, Russian Academy
of Sciences, Moscow, Russia

Vladimir A. Efimushkin
Central Research Institute of Communications, Moscow, Russia

Tatiana A. Gavrilo
Saint-Petersburg University, St. Petersburg, Russia

Hervé Glotin
University of Toulon, La Garde, France

Alexey O. Golosov
FORS Development Center, Moscow, Russia

Andrey Yu. Gribov
CyberPlat Company, Moscow, Russia

Alexander I. Gromoff
National Research University Higher School of Economics,
Moscow, Russia

Vladimir A. Gurvich
Rutgers, The State University of New Jersey, Rutgers, USA

Laurence Jacobs
University of Zurich, Zurich, Switzerland

Liliya A. Demidova
Ryazan State Radio Engineering University, Ryazan, Russia

Iosif E. Diskin
Russian Public Opinion Research Center, Moscow, Russia

Nikolay I. Ilyin
Federal Security Guard of the Russian Federation,
Moscow, Russia

Dmitry V. Isaev
National Research University Higher School of Economics,
Moscow, Russia

Alexander D. Ivannikov
Institute for Design Problems in Microelectronics, Russian Academy
of Sciences, Moscow, Russia

Valery A. Kalyagin
National Research University Higher School of Economics,
Nizhny Novgorod, Russia

Tatiana K. Kravchenko
National Research University Higher School of Economics,
Moscow, Russia

Sergei O. Kuznetsov
National Research University Higher School of Economics,
Moscow, Russia

Kwei-Jay Lin
Nagoya Institute of Technology, Nagoya, Japan

Mikhail I. Lugachev
Lomonosov Moscow State University, Moscow, Russia

Svetlana V. Maltseva
National Research University Higher School of Economics,
Moscow, Russia

Peter Major
UN Commission on Science and Technology for Development,
Geneva, Switzerland

Boris G. Mirkin
National Research University Higher School of Economics,
Moscow, Russia

Vadim V. Mottl
Tula State University, Tula, Russia

Dmitry M. Nazarov
Ural State University of Economics, Ekaterinburg, Russia

Dmitry E. Palchunov
Novosibirsk State University, Novosibirsk, Russia

Panagote (Panos) M. Pardalos
University of Florida, Gainesville, USA

Óscar Pastor
Polytechnic University of Valencia, Valencia, Spain

Joachim Posegga
University of Passau, Passau, Germany

Konstantin E. Samouylov
Peoples' Friendship University, Moscow, Russia

Kurt Sandkuhl
University of Rostock, Rostock, Germany

Yuriy D. Shmidt
Far Eastern Federal University, Vladivostok, Russia

Christine Strauss
University of Vienna, Vienna, Austria

Ali R. Sunyaev
Karlsruhe Institute of Technology, Karlsruhe, Germany

Victor V. Taratukhin
University of Munster, Munster, Germany

José M. Tribolet
Universidade de Lisboa, Lisbon, Portugal

Olga A. Tsukanova
Saint-Petersburg National Research University of Information
Technologies, Mechanics and Optics, St. Petersburg, Russia

Mikhail V. Ulyanov
V.A. Trapeznikov Institute of Control Sciences, Russian Academy
of Sciences, Moscow, Russia

Raissa K. Uskenbayeva
International Information Technology University, Almaty, Kazakhstan

Markus Westner
Regensburg University of Applied Sciences, Regensburg, Germany

[DOI: 10.17323/2587-814X.2021.2.7.20](https://doi.org/10.17323/2587-814X.2021.2.7.20)

Цифровой завод: методы дискретно-событийного моделирования и оптимизации производственных характеристик

В.Л. Макаров^a 

E-mail: makarov@cemi.rssi.ru

А.Р. Бахтизин^a 

E-mail: albert@cemi.rssi.ru

Г.Л. Бекларян^a 

E-mail: glbeklaryan@gmail.com

А.С. Акопов^{b,a} 

E-mail: aakopov@hse.ru

^a Центральный экономико-математический институт, Российская академия наук
Адрес: 117418, г. Москва, Нахимовский проспект, д. 47

^b Национальный исследовательский университет «Высшая школа экономики»
Адрес: 101000, г. Москва, ул. Мясницкая, д. 20

Аннотация

В статье представлен новый подход к разработке «цифрового двойника» производственного предприятия, на примере завода по выпуску телевизионной техники. Особенностью предлагаемого подхода является применение гибридных методов агентного и дискретно-событийного имитационного моделирования, с использованием которых реализована компьютерная модель сложного производственного процесса сборки готовой продукции (телевизоров) из поступающих комплектующих. Важнейшими требованиями, предъявляемыми к подобной системе, является интегрируемость всех ключевых звеньев цифрового завода: конвейерных линий, складов с комплектующими и готовой продукцией, сортировочно-конвейерной системы, сборочного цеха, отдела технического контроля, упаковочного цеха и других подразделений. Предлагаемая имитационная модель реализована в системе AnyLogic, поддерживающей возможности применения методов агентного и дискретно-событийного моделирования в рамках одной модели, а также использования встроенного генетического алгоритма для оптимизации основных параметров модели — наиболее важных производственных характеристик (например, времени сборки одного изделия или числа сотрудников, задействованных в процессах сборки, контроля качества и упаковки). С помощью разработанной модели проведены оптимизационные эксперименты при различных интенсивностях загрузки конвейерных линий комплектующими, различных ограничениях на трудовые ресурсы и др. Исследованы три сценария поведения производственной системы: отсутствие дефицита комплектующих с возможностью существенного увеличения привлекаемых трудовых ресурсов, дефицит комплектующих при сохранении спроса на конечную продукцию, наличие жестких ресурсных ограничений на число сотрудников, задействованных в процессах на фоне профицита комплектующих.

Ключевые слова: цифровой завод; производственные процессы; агентное моделирование; дискретно-событийное моделирование; генетический алгоритм; AnyLogic.

Цитирование: Макаров В.Л., Бахтизин А.Р., Бекларян Г.Л., Акопов А.С. Цифровой завод: методы дискретно-событийного моделирования и оптимизации производственных характеристик // Бизнес-информатика. 2021. Т. 15. № 2. С. 7–20. DOI: 10.17323/2587-814X.2021.2.7.20

Введение

В настоящее время для многих промышленных предприятий становится все более актуальной задача рационального управления производственными характеристиками в условиях динамично меняющейся внешней среды (например, резкого снижения объемов поставок комплектующих, существенного увеличения стоимости сырья, дефицита трудовых ресурсов). При этом сложность производственных и логистических процессов зачастую не позволяет предприятиям быстро адаптироваться за счет оперативного управления собственными ресурсами (например, уменьшать объем производства, не сокращая при этом персонал, или менять структуру производственного портфеля без существенных капиталовложений). Причиной возникающих проблем, как правило, является несбалансированное состояние различных звеньев производственной цепочки (например, дефицит производственных мощностей при существенной скорости оборачиваемости складских запасов). Поэтому требуется разработка систем поддержки принятия решений (СППР), основанных на построении «цифровых двойников» предприятий, с помощью которых можно определить наилучшие значения характеристик производственного процесса при различных сценарных условиях (например, при резком изменении спроса на конечную продукцию или при возникновении перебоев в поставках сырья и комплектующих).

Первые работы, посвященные проблемам рационального производственного планирования, относятся к середине XX века. Следует выделить наиболее важные работы по исследованию рациональной организации производственной деятельности с использованием методов системной динамики [1, 2]. Также хорошо известны разработки по применению теории систем массового обслуживания (СМО) [3, 4] для управления сложными производственными и логистическими процессами [5, 6]. Главной задачей подобных методов, развиваемых в рамках известного научного направления исследования операций

[7], является оптимизация характеристик цепочек создания конечной продукции, например, минимизация затрат, максимизация скорости оборачиваемости складских запасов, оптимизация количества и структуры задействованных трудовых ресурсов, обеспечение требуемого уровня качества и своевременности поставок готовой продукции [8].

В настоящее время для моделирования производственных и финансовых характеристик предприятия применяются методы системной динамики [9–11], агентного моделирования [12–14], дискретно-событийного моделирования [15, 16], эвристические методы оптимизации класса генетических алгоритмов [17–19] и некоторые другие. Например, в работе [10] представлена разработанная с использованием методов системной динамики имитационная модель вертикально-интегрированной нефтяной компании, включающей звенья нефтедобычи, нефтепереработки, транспортировки и сбыта. Преимуществом данной модели является возможность максимизации акционерной стоимости нефтяной компании при множественных ограничениях и с учетом влияния системы обратных связей, возникающих в производственно-сбытовой цепочке.

Следующим этапом в развитии подобных интеллектуальных производственных систем является создание так называемых «цифровых двойников» [20–24], которые построены на имитационных моделях, интегрированных с базами данных, хранилищами данных и оптимизационными модулями. Так, в работе [20] представлен «цифровой двойник» организации, оказывающей финансовые услуги, а в статье [21] предложена имитационная модель добывающего предприятия, реализованная в системе Powersim.

В данной статье развивается методология построения «цифровых двойников» производственного предприятия на основе методов дискретно-событийного и агентного моделирования, с реализацией в системе AnyLogic [25]. Предложена новая имитационная модель класса «цифровой завод» (на примере предприятия, осуществляющего сборку телевизионной техники) и выполнена оптимизация наиболее

важных параметров производственной системы по критерию значения накопленной валовой прибыли при различных сценарных условиях.

1. Концепция цифрового завода

В настоящее время телевизионная техника на большинстве российских заводов собирается из готовых комплектующих, производство которых осуществляется на других предприятиях с использованием импортной электронной компонентой базы (например, микропроцессоров, ЖК-дисплеев и т.д.). Несмотря на определенные недостатки (такие как полная зависимость от поставок импортных комплектующих), подобная организация производственного процесса позволяет существенно сократить операционные издержки, в основном за счет привлечения менее дорогостоящей (местной) рабочей силы, а также минимизации транспортных и налоговых издержек.

Концептуальная модель «цифрового завода» основана на создании ряда взаимодействующих подсистем, среди которых наиболее важными являются следующие:

♦ **подсистема имитационного моделирования**, обеспечивающая возможность вычисления производственных и финансовых характеристик предприятия, с учетом детализации всей цепочки создания конечного продукта для управления на микроуровне. Имитационная модель предприятия, осуществляющего сборку телевизионной техники, включает следующие элементы:

1. **источники поставок комплектующих**, необходимых для сборки конечной продукции (в частности, к подобным комплектующим относятся корпуса, жидкокристаллические экраны, системные платы, видеоадаптеры, блоки обработки изображения, блоки питания и др.), поступающих на завод в соответствии с заданной интенсивностью;
2. **склады комплектующих**, предназначенные для временного хранения промежуточной продукции (деталей) и их поставок на конвейерные линии;
3. **сортировочно-конвейерная система**, предназначенная для автоматической синхронизации потоков с комплектующими, имеющими различные технические и эксплуатационные характеристики. К таким характеристикам, в частности, относятся размер жидкокристаллического экрана (который варьируется от 20 до 90 дюймов), а также размер корпуса;

4. **сборочный цех**, обеспечивающий сборку готовой продукции с использованием привлекаемых специалистов (трудовых ресурсов);
5. **отдел технического контроля**, обеспечивающий контроль качества конечной продукции (тестирование устройств);
6. **упаковочный цех**, выполняющий функцию упаковки, комплектования и формирования партий готовой продукции;
7. **склад готовой продукции**, предназначенный для временного хранения телевизионной техники;
8. **отдел отгрузки**, обеспечивающий функцию отгрузки партий готовой продукции посредством ее передачи транспортной компании;

♦ **хранилище данных**, предназначенное для хранения исходных данных, необходимых для имитационного моделирования и обработки результатов оптимизационных экспериментов;

♦ **оптимизационный модуль**, агрегированный по целевым функционалам с имитационной моделью предприятия и обеспечивающий возможность поиска наилучших значений характеристик производственного процесса при различных сценарных условиях.

Укрупненная схема «цифрового двойника» предприятия, осуществляющего сборку телевизионной техники, представлена на *рисунке 1*.

Отметим, что важной особенностью рассматриваемой системы (*рисунок 1*) является наличие агентов, участвующих в производственном процессе. К ним относятся:

- ♦ **агенты-комплектующие** со своими индивидуальными характеристиками;
- ♦ **агенты-изделия**, дифференцируемые по своим функциональным характеристикам (например, модель телевизора, размер матрицы, стоимость и т.д.);
- ♦ **агенты-партии** готовой продукции, представляющие собой группировку агентов-изделий (по моделям) для транспортировки;
- ♦ **агенты-ресурсы**, задействованные в различных процессах (например, сборщики, специалисты ОТК, упаковщики и др.).

При этом производственная система обеспечивает механизм автоматической трансформации агентов-комплектующих в агентов-готовые изделия, а также группировку агентов-изделий в партии готовой продукции.

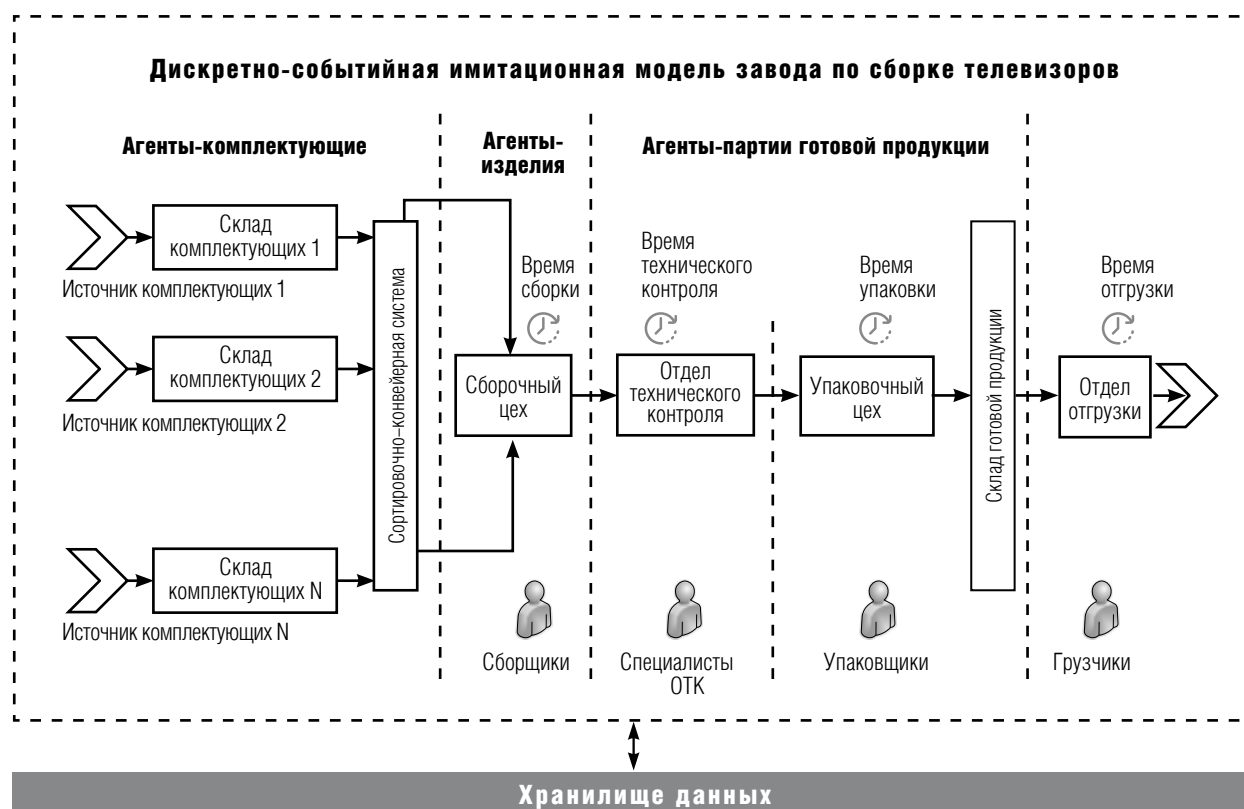


Рис. 1. Укрупненная схема «цифрового двойника» производственного предприятия

2. Имитационная модель

Ранее, в работе [21] была представлена имитационная модель сборочного цеха, реализованная в системе AnyLogic. Эта модель может рассматриваться в качестве прототипа при создании более сложной модели гибридного типа, в которой одновременно применяются методы системной динамики, агентного и дискретно-событийного моделирования. Другим важным отличием предлагаемой в данной статье модели является более сложная структура и динамика поставок комплектующих, характеризующаяся наличием множественных временных разрывов. При этом основные сервисные характеристики производственного процесса (в частности, время сборки, технического контроля, отгрузки и т.д.), существенно влияющие на его результативность, являются стохастическими. В результате метрика результативности подобной системы (динамика прибыли) относится к классу мультимодальных целевых функций с множественными точками разрыва, максимизация которой не может быть выполнена известными аналитическими методами. Поэтому для создания «цифрового двойника» производственного предприятия, осу-

ществляющего сборку телевизионной техники, разрабатывается оригинальная имитационная модель, агрегированная по целевому функционалу со встроенным генетическим алгоритмом, обеспечивающим поиск наилучших решений на основе эвристики.

Приведем краткую постановку задачи оптимизации ключевых характеристик рассматриваемого производственного предприятия. Для этого введем следующие обозначения:

♦ $t_r \in T$ – дискретное модельное время; $T = \{t_1, t_2, \dots, t_{|T|}\}$ – набор моментов модельного времени; $|T|$ – общее число моментов модельного времени;

♦ $I = \{i_1, i_2, \dots, i_{|I|}\}$ – набор индексов моделей собираемой телевизионной техники (дифференцируемой, в частности, по размеру экрана, функциональным характеристикам и др.); $|I|$ – общее число собираемых моделей телевизоров;

♦ $J_i = \{j_1, j_2, \dots, j_{|J_i|}\}, i \in I$ – набор индексов комплектующих, необходимых для сборки телевизионной техники i -го типа (например, корпуса, ЖК экраны, системные платы и др.); $|J_i|$ – общее число

комплектующих, необходимых для сборки телевизора i -го типа;

♦ $\{n_1, n_2, \dots, n_N\}$ – число специалистов различных категорий (сборщиков первой и второй категорий, специалистов ОТК, упаковщиков и грузчиков), задействованных в производственных процессах, чел.; N – общее количество квалификационных категорий сотрудников; $\xi = \{1, 2, \dots, N\}$ – индекс категории сотрудников и типов процессов, в которых они задействованы;

♦ $\{w_1(t_r), w_2(t_r), \dots, w_N(t_r)\}, t_r \in T$ – суточная заработная плата специалистов различных категорий, задействованных в производственных процессах, руб.;

♦ $c_{j_i}(t_r), j_i \in J_i, i \in I, t_r \in T$ – стоимость комплектующих, требуемых для сборки i -й модели телевизора в момент времени t_r , руб.;

♦ $p_i(t_r), i \in I, t_r \in T$ – стоимость единицы готовой продукции (телевизора) в момент времени t_r , руб.;

♦ $\{\tau_1, \tau_2, \dots, \tau_N\}$ – среднее время сборки готового изделия из основных и вспомогательных деталей, время технического контроля, время упаковки и отгрузки в расчете на одного сотрудника, задействованного в соответствующих операциях;

♦ $\delta_{j_i}(t_r), j_i \in J_i, i \in I, t_r \in T$ – эндогенная структура поставок комплектующих в момент времени t_r , задаваемая с использованием единичной импульсной функции со случайными аргументами, шт.;

♦ $v_i(t_r), i \in I, t_r \in T$ – эндогенная структура выпуска готовой продукции в момент времени t_r , определяемая в результате имитационного моделирования и зависящая от наборов управляющих параметров $\{n_1, n_2, \dots, n_N\}, \{\tau_1, \tau_2, \dots, \tau_N\}$, а также динамики поставок комплектующих $\delta_{j_i}(t_r)$, шт.

Суточная валовая прибыль предприятия в момент времени $t_r, t_r \in T$:

$$\pi(t_r) = \sum_{i=1}^{|I|} \left((v_i(t_r) p_i)(t_r) - \sum_{j_i=1}^{|J_i|} \delta_{j_i}(t_r) c_{j_i}(t_r) \right) - \sum_{\xi=1}^N w_{\xi}(t_r) n_{\xi}. \quad (1)$$

Теперь можно сформулировать центральную задачу производственного предприятия по максимизации накопленной валовой прибыли.

Задача. Требуется максимизировать накопленную валовую прибыль по наборам управляющих параметров $\{n_1, n_2, \dots, n_N\}$ и $\{\tau_1, \tau_2, \dots, \tau_N\}$:

$$\sum_{t_r=1}^{|T|} \pi(t_r) \rightarrow \max_{\{n_1, n_2, \dots, n_N\}, \{\tau_1, \tau_2, \dots, \tau_N\}}, \quad (2)$$

при ограничениях: $n_{\xi} \leq n_{\xi} \leq \bar{n}_{\xi}, \tau_{\xi} \leq \tau_{\xi} \leq \bar{\tau}_{\xi}$.

Здесь $\{n_{\xi}, \bar{n}_{\xi}\}, \{\tau_{\xi}, \bar{\tau}_{\xi}\}$ – известные граничные (верхние и нижние) значения управляющих параметров.

Предложенная имитационная модель, обеспечивающая вычисление целевого функционала (1)–(2), реализована в системе AnyLogic (рисунк 2). Отличительной особенностью модели является высокая размерность потоков исходных данных (десятки тысяч комплектующих, участвующих в процессе сборки телевизионной техники со своими индивидуальными характеристиками), что, в частности, обуславливается наличием большого числа элементов типа «склад», используемых для хранения промежуточной и конечной продукции, а также применением сортировочно-конвейерных систем для синхронизации потоков агентов-комплектующих (рисунк 2).

Другой важной особенностью модели, представленной на рисунке 2, является интеграция блока расчета финансово-экономических показателей с операционными подмоделями, реализующими процессы получения и распределения комплектующих посредством сортировочно-конвейерных систем, сборку, технический контроль, упаковку и формирование партий, дифференцируемых по видам продукции (например, в зависимости от размера диагонали экрана и заданного набора функций). Исходные данные по агентам-комплектующим (характеристики деталей) загружаются из базы данных системы с учетом переменного числа деталей каждого вида, одновременно поступающих в соответствии с заданной интенсивностью. Программная реализация имитационной модели выполнена с использованием языка программирования Java, поддерживаемого в системе AnyLogic.

Для запуска последовательной обработки комплектующих в модели используется специальный объект типа **source** (источник комплектующих), в котором выполняется переопределение характеристик каждого генерируемого агента, принадлежащего к классу **Components_X** (комплектующие типа **X**), в специальном поле – «Действия при подходе к выходу»:

```
((Components_X) entity).parameter = components_X.  
get(index_X).parameter;  
index_X = index_X ++;
```

Здесь **Components_X** – популяция агентов-комплектующих типа **X** с характеристиками, значения которых динамически загружаются из хранилища данных (MS SQL Server), **parameter** – один

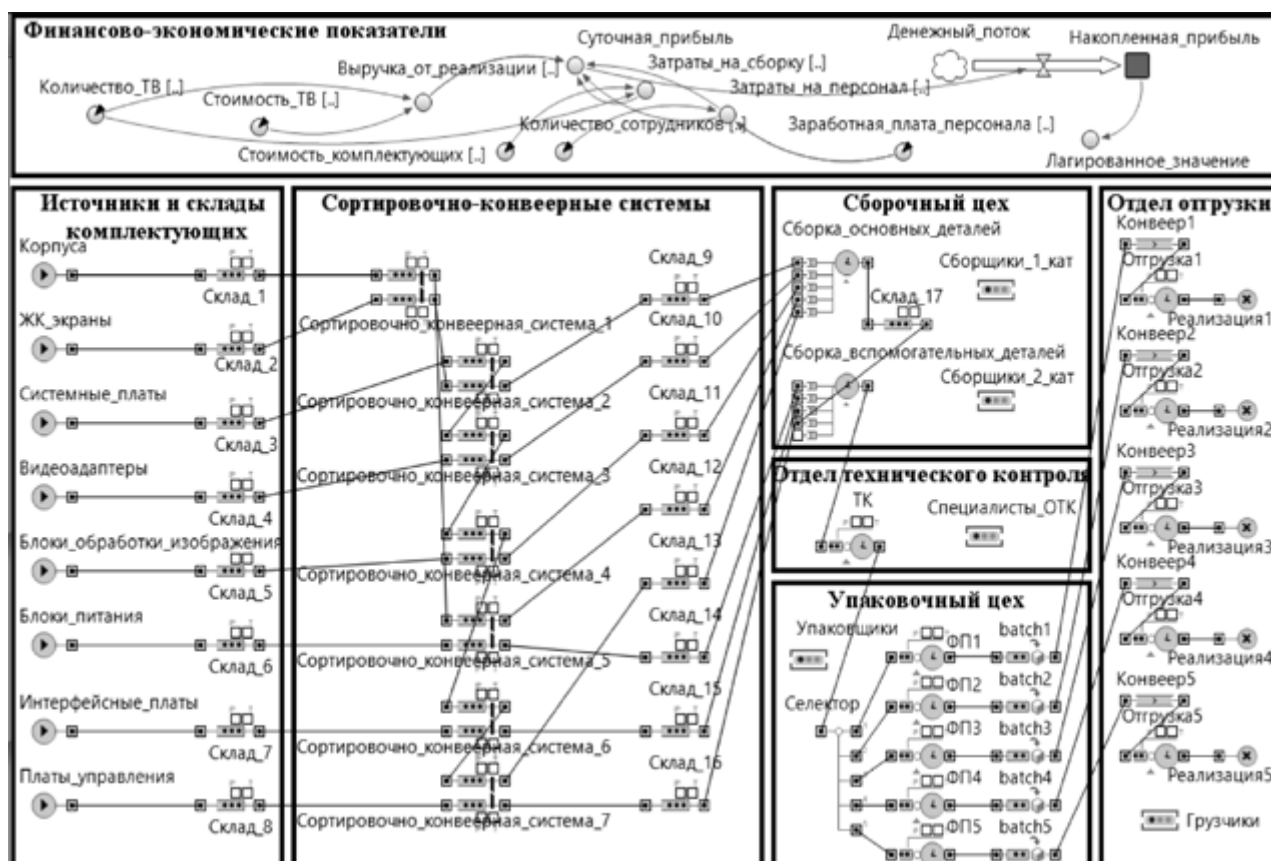


Рис. 2. Имитационная модель завода по производству телевизионной техники в AnyLogic

возможных параметров агента-комплектующего ((**Components_X**) entity), например, габаритный размер корпуса, **index_X** — глобальная переменная, используемая как итератор для доступа к очередному члену популяции **Components_X** (т.е. агенту-комплектующему). Отметим, что, начиная с версии AnyLogic 8, в системе поддерживается более простая форма доступа к данным агента, участвующего в процессной цепочке, с использованием атрибута **agent** вместо ((**Components_X**) entity). При этом при создании множественных заявок метод «**Действия при подходе к выходу**» также выполняется многократно для каждого агента-комплектующего, исходящего из источника (рисунки 3). Количество агентов-комплектующих генерируемых в каждый момент времени, задается посредством единичной импульсной функции со случайными аргументами, определяющими вероятное количество деталей, которые поступят в случае прихода данной партии:

```
(int) pulse(uniform(0, 100), uniform(0, 100))*uniform_discr(1000, 1500);
```

Также одним из наиболее важных элементов предлагаемой модели является формирование партий готовой продукции посредством распределения потока прошедшей технический контроль телевизионной техники на группы в зависимости от типа агента-изделия (т.е. модели телевизора) (рисунки 4).

Поскольку стоимость готовой продукции зависит от модельного ряда собираемой телевизионной техники, при расчете выручки и затрат можно использовать показатели, значения которых вычисляются в элементах диаграммы процессов (рисунки 2). Например, количество собираемых телевизоров различных типов задается следующим образом:

```
Количество_TB.set(Реализация1.count() * batch1_size, M);
```

Здесь **Реализация1.count()** — функция, обеспечивающая вычисление количества элементов (агентов-изделий первого вида) на выходе из производственного процесса, **batch1_size** — количество изделий в партии, **M** — индекс модели телевизора.

Корпуса - Source

Имя: ☒ Отображать имя ☐ Исключить

Тип заявки:

Прибывают согласно:

Интенсивность прибытия:

За 1 раз создается несколько заявок: ☒

Кол-во заявок, прибывающих за 1 раз:

Ограниченное кол-во прибытий: ☐

Новая заявка: (чтобы создать другой тип, добавьте его из палитры)

Местоположение прибытия:

Специфические

Установить время начала: ☐

Добавить заявки в: ☒ Популяцию по умолчанию ☐ Другую популяцию агентов

Выталкивать заявки: ☒

Действия

До прибытия:

При подходе к выходу:

```
((Components_1) entity).parameter1 = components_1.get(index_1).parameter1;
((Components_1) entity).parameter2 = components_1.get(index_1).parameter2;
((Components_1) entity).parameter3 = components_1.get(index_1).parameter3;
((Components_1) entity).parameter4 = components_1.get(index_1).parameter4;
((Components_1) entity).parameter5 = components_1.get(index_1).parameter5;
index_1 = index_1 + 1;
```

При выходе:

Специфические

Описание

Рис. 3. Свойства типового источника поставок агентов-комплектующих

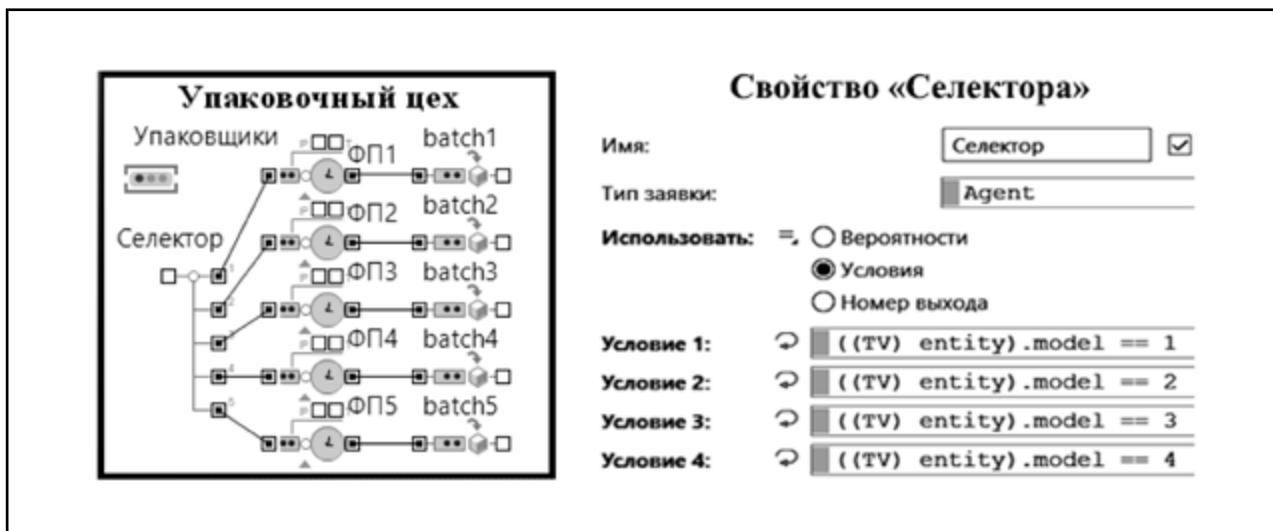


Рис. 4. Распределение потока готовой продукции при формировании партий

3. Результаты экспериментов

Как было отмечено ранее, источники поставок агентов-комплектующих обеспечивают генерацию необходимых для сборки компонент посредством единичной импульсной функции со случайными аргументами. На *рисунке 5* показана динамика подобных поставок на интервале модельного времени от 1 до 100 дней.

Как следует из *рисунка 5*, между поставками комплектующих существуют временные разрывы, обусловленные естественными логистическими сбоями. Поэтому в модели обеспечивается автоматическая синхронизация входных потоков посредством сортировочно-конвейерных систем, перенаправляющих отобранные и взаимно соответствующие детали в сборочный цех.

На *рисунке 6* представлена модельная динамика складских запасов основных и вспомогательных деталей, а также готовых изделий (телевизоров).

Далее были выполнены оптимизационные эксперименты, направленные на максимизацию прибыли, формируемой в результате сборки и реализации телевизионной техники. Оптимизационные эксперименты были выполнены с использованием встроенного в AnyLogic генетического алгоритма. При этом в качестве переменных-решений выбраны следующие управляющие параметры модели:

♦ **количество агентов-сборщиков** (людей) различных категорий, участвующих в процессах сборки основных и вспомогательных деталей;

♦ **количество агентов-специалистов ОТК**, отвечающих за выполнение технического контроля готовой продукции;

♦ **количество агентов-упаковщиков и агентов-грузчиков**, участвующих в процессах упаковки и отгрузки готовой продукции;

♦ **средние (медианные) значения времени сборки, технического контроля, упаковки и отгрузки** (реализации), влияющие на скорость исполнения соответствующих процессов.

При этом рассматривались три сценария поведения моделируемой производственной системы:

Сценарий 1. Отсутствие дефицита комплектующих с возможностью существенного увеличения привлекаемых трудовых ресурсов.

Сценарий 2. Дефицит комплектующих при сохранении спроса на конечную продукцию.

Сценарий 3. Наличие жестких ресурсных ограничений на число сотрудников, задействованных в процессах на фоне профицита комплектующих.

На *рисунке 7* представлены результаты оптимизационных экспериментов.

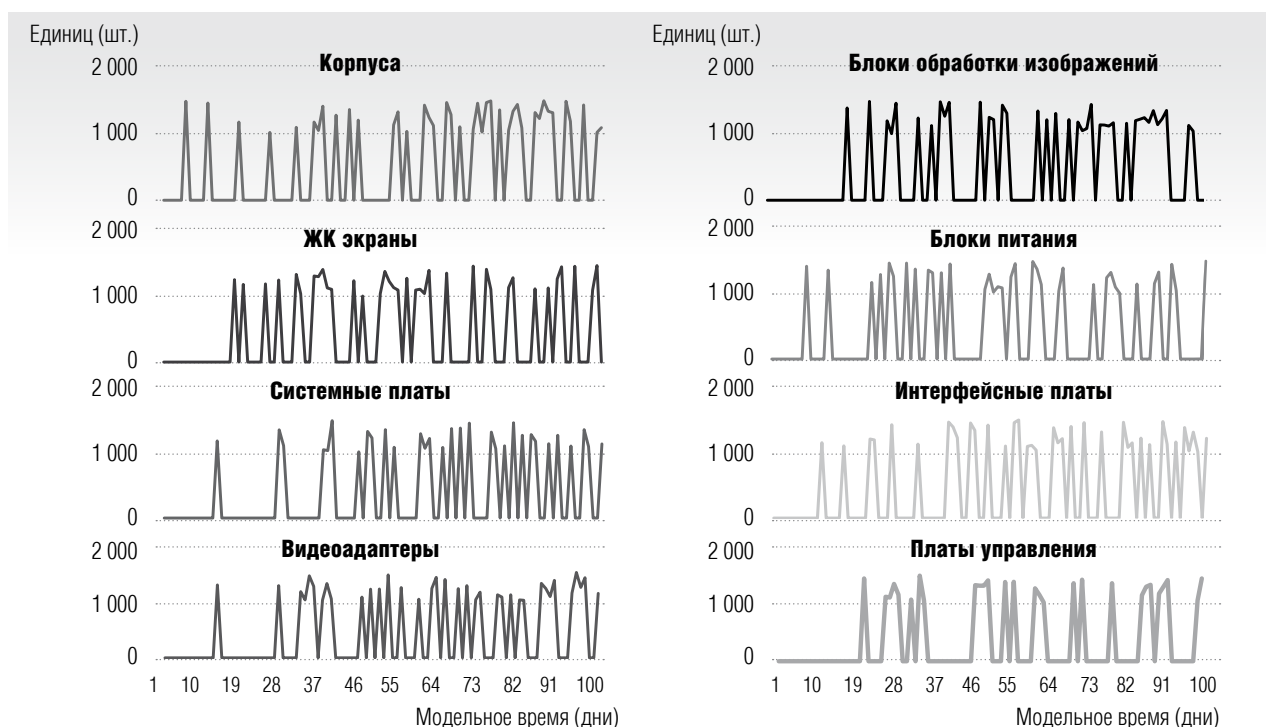


Рис. 5. Модельная динамика поставки комплектующих

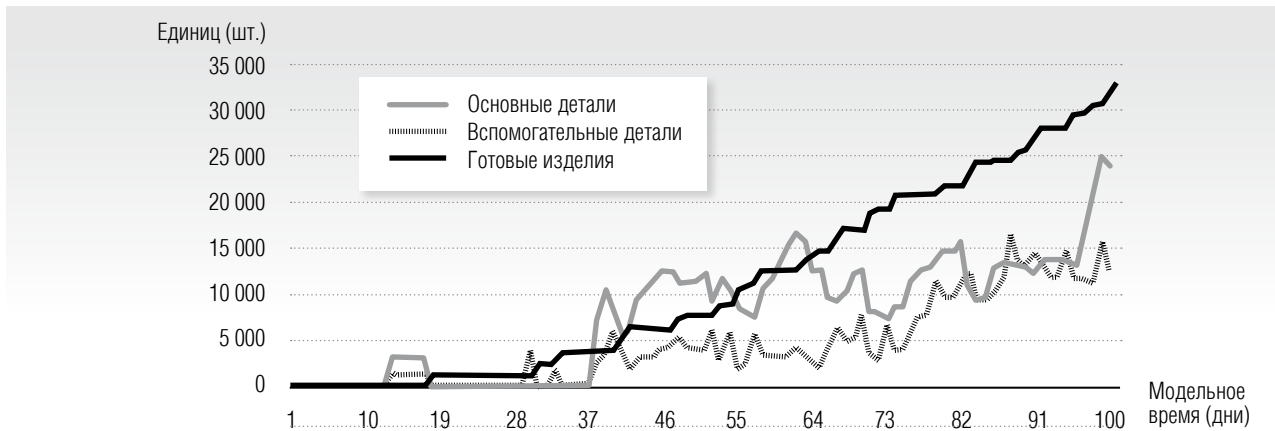


Рис. 6. Модельная динамика складских запасов

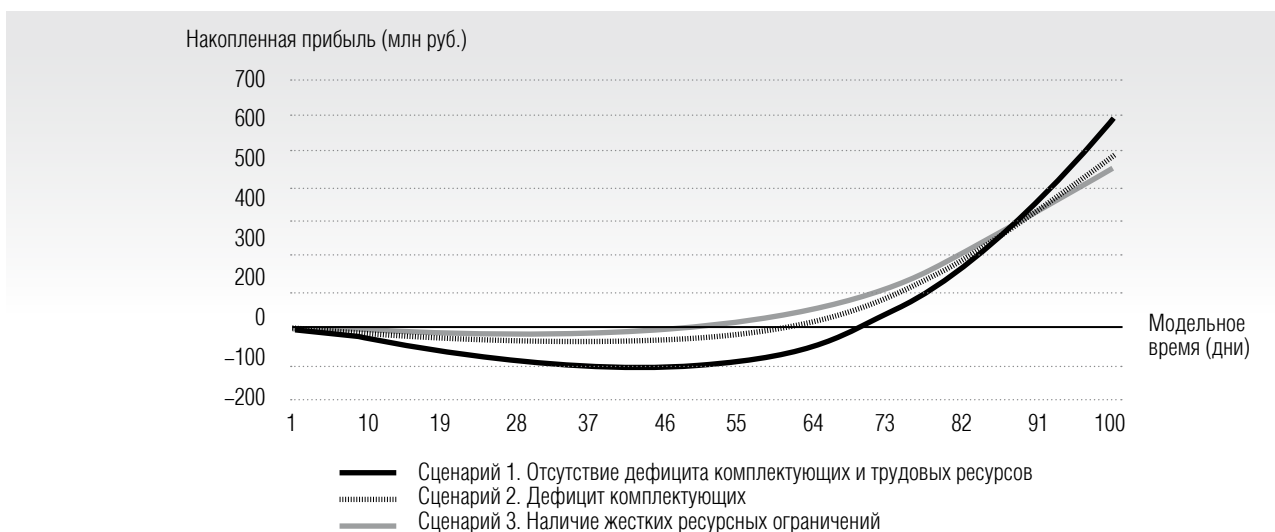


Рис. 7. Результаты оптимизационных экспериментов (накопленная прибыль)

В таблице 1 показаны численные значения искомым переменных, полученные в результате оптимизации для трех рассматриваемых сценариев поведения исследуемой производственной системы.

Как следует из рисунка 7 и таблицы 1, применение генетического алгоритма для оптимизации характеристик разработанной дискретно-событийной имитационной модели позволило определить наилучшие (субоптимальные) значения искомым переменных, соответствующие трем принципиально различающимся сценариям поведения производственной системы. При этом динамика накопленной валовой прибыли в основном устойчива по отношению к исследуемым сценариям. Например, в условиях дефицита комплектующих (Сценарий 2) за счет значительного сокращения количества при-

влеченных трудовых ресурсов при одновременном уменьшении времени выполнения производственных операций (т.е. при увеличении производительности труда) можно обеспечить значительный уровень накопленной прибыли. Также близкий финансовый результат можно получить в условиях жестких ресурсных ограничений и отсутствия дефицита комплектующих (Сценарий 3).

Заключение

В данной статье представлен новый подход к разработке «цифрового двойника» производственного предприятия (на примере завода по сборке телевизионной техники), основанный на комбинированном применении методов дискретно-событийного и агентного имитационного моделирования. Пред-

Таблица 1.

Значения искомых переменных и результаты оптимизации

Параметр модели	Сценарий 1: Отсутствие дефицита комплектующих и трудовых ресурсов	Сценарий 2: Дефицит комплектующих	Сценарий 3: Наличие жестких ресурсных ограничений
Количество агентов–сборщиков первой категории, чел.	1000	430	125
Количество агентов–сборщиков второй категории, чел.	500	225	84
Количество агентов–специалистов ОТК, чел.	100	30	15
Количество агентов–упаковщиков, чел.	100	27	13
Количество агентов–грузчиков, чел.	50	20	10
Среднее время сборки, дней	0,1	0,05	0,03
Среднее технического контроля, дней	0,01	0,01	0,008
Среднее время упаковки, дней	0,01	0,07	0,05
Среднее время отгрузки, дней	7	5	4
Накопленная валовая прибыль в момент времени T=100, млн. руб.	619	494	468

ложена концептуальная модель цифрового завода, с последующей программной реализацией в системе имитационного моделирования AnyLogic. Сформулирована и решена важнейшая задача производственного предприятия по максимизации валовой накопленной прибыли с учетом вклада операционной составляющей, в том числе динамики объема выпуска готовой продукции, численности сотрудников, задействованных в различных процессах, времени исполнения производственных операций и др. Выполнены численные исследования, в частности, оптимизационные эксперименты, подтверждающие возможность определения наилучших решений, обеспечивающих устойчивость накопленной валовой прибыли предприятия по отношению к различным характеристикам внешней среды (например, в условиях дефицита комплектующих и жестких ресурсных ограничений).

Результаты имитационного моделирования демонстрируют возможность получения накопленной валовой прибыли производственного предприятия,

устойчивой по отношению к сценарным характеристикам внешней среды. Например, при возникновении дефицита комплектующих (Сценарий 2), а также в условиях жестких ресурсных ограничений и отсутствия дефицита комплектующих (Сценарий 3) можно обеспечить уровень накопленной прибыли близкой к финансовому результату, характерному для первого сценария с отсутствием подобных ограничений.

Дальнейшие исследования будут направлены на развитие методов разработки «цифровых двойников» производственных предприятий с более сложной организационной структурой и вложенными (иерархическими, многоуровневыми) процессами. ■

Благодарности

Исследование выполнено при финансовой поддержке Российского фонда фундаментальных исследований (РФФИ), в рамках научного проекта № 18-29-03139.

Литература

1. Forrester J.W. Industrial dynamics. MIT Press, 1961.
2. Sterman J.D. Business dynamics: Systems thinking and modeling for a complex world. McGraw Hill, 2000.
3. Saaty T.L. Elements of queuing theory with applications. Dover Pubns, 1983.

4. Sundarapandian V. Probability, statistics and queueing theory. PHI Learning, 2009.
5. Rashid R., Hoseini S.F., Gholamian M.R., Feizabadi M. Application of queueing theory in production-inventory optimization // *International Journal of Industrial Engineering*. 2015. No 11. P. 485–494. DOI: 10.1007/s40092-015-0115-9.
6. Marsudi M., Shafeek H. Production line performance by using queueing model // 7th IFAC Conference on Manufacturing Modelling, Management, and Control. Saint Petersburg, Russia, 19–21 June 2013. Vol. 46. No 9. P. 1152–1157.
7. Wetherbe J.C. Systems analysis for computer-based information systems. West Publishing, 1979.
8. Robinson S. Simulation: The practice of model development and use. Wiley, 2004.
9. Акопов А.С., Хачатрян Н.К. Системная динамика. М.: ЦЭМИ РАН, 2014.
10. Akopov A.S. Designing of integrated system-dynamics models for an oil company // *International Journal of Computer Applications in Technology*. 2012. Vol. 45. No. 4. P. 220–230. DOI: 10.1504/IJCAT.2012.051122.
11. Beklaryan G.L., Akopov A.S., Khachatryan N.K. Optimisation of system dynamics models using a real-coded genetic algorithm with fuzzy control // *Cybernetics and Information Technologies*. 2019. Vol. 19. No. 2. P. 87–103. DOI: 10.2478/cait-2019-0017.
12. Акопов А.С., Хачатрян Н.К. Агентное моделирование. М.: ЦЭМИ РАН, 2016.
13. Бахтизин А.Р. Агент-ориентированные модели экономики. М.: Экономика, 2008.
14. Akopov A.S., Beklaryan L.A., Saghatelian A.K. Agent-based modelling of interactions between air pollutants and greenery using a case study of Yerevan, Armenia // *Environmental Modelling and Software*. 2019. No 116. P. 7–25. DOI: 10.1016/j.envsoft.2019.02.003.
15. Акопов А.С. Имитационное моделирование. М.: Юрайт, 2014.
16. Tang J., Leu G., Abbass H.A. Discrete event simulation / Simulation and Computational Red Teaming for Problem Solving. IEEE, 2020. P. 121–142.
17. Akopov A.S., Beklaryan L.A., Beklaryan A.L. Cluster-based optimization of an evacuation process using a parallel bi-objective real-coded genetic algorithm // *Cybernetics and Information Technologies*. 2020. Vol. 20. No. 3. P. 45–63. DOI: 10.2478/cait-2020-0027.
18. Akopov A.S., Beklaryan L.A., Thakur M., Verma D.B. Parallel multi-agent real-coded genetic algorithm for large-scale black-box single-objective optimization // *Knowledge-Based Systems*. 2019. No 174. P. 103–122. DOI: 10.1016/j.knsys.2019.03.003.
19. Akopov A.S. Parallel genetic algorithm with fading selection // *International Journal of Computer Applications in Technology*. 2014. Vol. 49. No. 3/4. P. 325–331. DOI: 10.1504/IJCAT.2014.062368.
20. Макаров В.Л., Бахтизин А.Р., Бекларян Г.Л., Акопов А.С. Разработка цифровых двойников для финансовых организаций // *Аудит и финансовый анализ*. 2019. № 5. С. 38–45.
21. Makarov V.L., Bakhtizin A.R., Beklaryan G.L. Developing digital twins for production enterprises // *Business Informatics*. 2019. Vol. 14. No 1. P. 7–16. DOI: 10.17323/1998-0663.2019.4.7.16.
22. Saddik A.El. Digital twins: The convergence of multimedia technologies // *IEEE MultiMedia*. 2018. Vol. 25. No 2. P. 87–92. DOI: 10.1109/MMUL.2018.023121167.
23. Digital twin-driven product design framework / F. Tao [et al.] // *International Journal of Production Research*. 2018. Vol. 57. No 12. P. 3935–3953. DOI: 10.1080/00207543.2018.1443229.
24. Xu Y., Sun Y., Liu X., Zheng Y. A digital-twin-assisted fault diagnosis using deep transfer learning // *IEEE Access*. 2019. No 7. P. 19990–19999. DOI: 10.1109/ACCESS.2018.2890566.
25. Borshchev A. The Big Book of simulation modeling. Multimethod modeling with AnyLogic 6. AnyLogic North America, 2013.

Об авторах

Макаров Валерий Леонидович

доктор физико-математических наук; академик Российской академии наук;
 научный руководитель, Центральный экономико-математический институт, Российская академия наук, 117418, г. Москва, Нахимовский проспект, д. 47;
 E-mail: makarov@cemi.rssi.ru
 ORCID: 0000-0002-2802-2100

Бахтизин Альберт Рауфович

доктор экономических наук; член-корреспондент Российской академии наук;
 директор, Центральный экономико-математический институт, Российская академия наук, 117418, г. Москва, Нахимовский проспект, д. 47;
 E-mail: albert@cemi.rssi.ru
 ORCID: 0000-0002-9649-0168

Бекларян Гаянэ Леоновна

кандидат экономических наук;
 старший научный сотрудник, лаборатория компьютерного моделирования социально-экономических процессов, Центральный экономико-математический институт, Российская академия наук, 117418, г. Москва, Нахимовский проспект, д. 47;
 E-mail: glbeklaryan@gmail.com
 ORCID: 0000-0002-1286-0345

Акопов Андраник Сумбатович

доктор технических наук;

профессор, департамент бизнес-информатики, Высшая школа бизнеса, Национальный исследовательский университет «Высшая школа экономики», 101000, г. Москва, ул. Мясницкая, д. 20;

главный научный сотрудник, лаборатория динамических моделей экономики и оптимизации, Центральный экономико-математический институт, Российская академия наук, 117418, г. Москва, Нахимовский проспект, д. 47;

E-mail: aakopov@hse.ru

ORCID: 0000-0003-0627-3037

Digital plant: methods of discrete-event modeling and optimization of production characteristics

Valery L. Makarov^a

E-mail: makarov@cemi.rssi.ru

Albert R. Bakhtizin^a

E-mail: albert@cemi.rssi.ru

Gayane L. Beklaryan^a

E-mail: glbeklaryan@gmail.com

Andranik S. Akopov^{b,a}

E-mail: aakopov@hse.ru

^a Central Economics and Mathematics Institute, Russian Academy of Sciences
Address: 47, Nakhimovsky Prospect, Moscow 117418, Russia

^b National Research University Higher School of Economics
Address: 20, Myasnitskaya Street, Moscow 101000, Russia

Abstract

This article presents a new approach to the development of a ‘digital twin’ of a manufacturing enterprise, using a television manufacturing plant as the case study. The feature of the proposed approach is the use of hybrid methods of agent-based modeling and discrete-event simulation in order to implement a simulation model of a complex production process for assembling final products from supplied components. The most important requirement for such a system is the integration of all key chains of a digital plant: conveyor lines, warehouses with components and final products (TVs), sorting and conveyor system, assembly unit, technical control department, packing unit, etc. The proposed simulation model is implemented in the AnyLogic system, which supports the possibility of using agent-based and discrete-event modeling methods within one model. The system also supports using the built-in genetic algorithm to optimize the main parameters of the model: the most important production characteristics (for example, assembly time of a product, the number of employees involved in assembly, quality control and packaging processes). Optimization experiments were completed with the help of the developed model at various intensities of loading conveyor lines with components, various restrictions on labor resources, etc. Three scenarios of the production system behavior are investigated: the absence of the components deficit with the possibility of significantly increasing the labor resource involved, a components deficit while demand for final products is maintained, and the presence of hard restrictions on the number of employees who can be involved in the processes under conditions of components deficit.

Key words: digital plant; production processes; agent-based modeling; discrete-event simulation; genetic algorithm; AnyLogic.

Citation: Makarov V.L., Bakhtizin A.R., Beklaryan G.L., Akopov A.S. (2021) Digital plant: methods of discrete-event modeling and optimization of production characteristics. *Business Informatics*, vol. 15, no 2, pp. 7–20.
DOI: 10.17323/2587-814X.2021.2.7.20

References

1. Forrester J.W. (1961) *Industrial dynamics*. MIT Press.
2. Sterman J.D. (2000) *Business dynamics: Systems thinking and modeling for a complex world*. McGraw Hill.
3. Saaty T.L. (1983) *Elements of queueing theory with applications*. Dover Pubns.
4. Sundarapandian V. (2009) *Probability, statistics and queueing theory*. PHI Learning.
5. Rashid R., Hoseini S.F., Gholamian M.R., Feizabadi M. (2015) Application of queueing theory in production-inventory optimization. *International Journal of Industrial Engineering*, no 11, pp. 485–494. DOI: 10.1007/s40092-015-0115-9.
6. Marsudi M., Shafeek H. (2013) Production line performance by using queueing model. *Proceedings of the 7th IFAC Conference on Manufacturing Modelling, Management, and Control, Saint Petersburg, Russia, 19–21 June 2013*, vol. 46, no 9, pp. 1152–1157.
7. Wetherbe J.C. (1979) *Systems analysis for computer-based information systems*. West Publishing.
8. Robinson S. (2004) *Simulation: The practice of model development and use*. Wiley.
9. Akopov A.S., Khachatryan N.K. (2014) *System dynamics*. Moscow: CEMI RAS (in Russian).
10. Akopov A.S. (2012) Designing of integrated system-dynamics models for an oil company. *International Journal of Computer Applications in Technology*, vol. 45, no 4, pp. 220–230. DOI: 10.1504/IJCAT.2012.051122.
11. Beklaryan G.L., Akopov A.S., Khachatryan N.K. (2019) Optimisation of system dynamics models using a real-coded genetic algorithm with fuzzy control. *Cybernetics and Information Technologies*, vol. 19, no 2, pp. 87–103. DOI: 10.2478/cait-2019-0017.
12. Akopov A.S., Khachatryan N.K. (2016) *Agent-based modeling*. Moscow: CEMI RAS (in Russian).
13. Bakhtizin A.R. (2008) *Agent-based models of the economy*. Moscow: Economics (in Russian).
14. Akopov A.S., Beklaryan L.A., Saghatelian A.K. (2019) Agent-based modelling of interactions between air pollutants and greenery using a case study of Yerevan, Armenia. *Environmental Modelling and Software*, no 116, pp. 7–25. DOI: 10.1016/j.envsoft.2019.02.003.
15. Akopov A.S. (2017) *Simulation modeling*. Moscow: Urait (in Russian).
16. Tang J., Leu G., Abbass H.A. (2020) Discrete event simulation. *Simulation and Computational Red Teaming for Problem Solving*. IEEE, pp. 121–142.
17. Akopov A.S., Beklaryan L.A., Beklaryan A.L. (2020) Cluster-based optimization of an evacuation process using a parallel bi-objective real-coded genetic algorithm. *Cybernetics and Information Technologies*, vol. 20, no 3, pp. 45–63. DOI: 10.2478/cait-2020-0027.
18. Akopov A.S., Beklaryan L.A., Thakur M., Verma D.B. (2019) Parallel multi-agent real-coded genetic algorithm for large-scale black-box single-objective optimization. *Knowledge-Based Systems*, no 174, pp. 103–122. DOI: 10.1016/j.knsys.2019.03.003.
19. Akopov A.S. (2014) Parallel genetic algorithm with fading selection. *International Journal of Computer Applications in Technology*, vol. 49, no 3/4, pp. 325–331. DOI: 10.1504/IJCAT.2014.062368.
20. Makarov V.L., Bakhtizin A.R., Beklaryan G.L., Akopov A.S. (2019) Developing digital twins for financial organisations. *Audit and Financial Analysis*, no 5, pp. 38–45 (in Russian).
21. Makarov V.L., Bakhtizin A.R., Beklaryan G.L. (2019) Developing digital twins for production enterprises. *Business Informatics*, vol. 14, no 1, pp. 7–16. DOI: 10.17323/1998-0663.2019.4.7.16.
22. Saddik A.El. (2018) Digital twins: The convergence of multimedia technologies. *IEEE MultiMedia*, vol. 25, no 2, pp. 87–92. DOI: 10.1109/MMUL.2018.023121167.
23. Tao F., Sui F., Liu A., Qi Q., Zhang M., Song B., Guo Z., Lu S.C.-Y., Nee A.Y.C. (2018) Digital twin-driven product design framework. *International Journal of Production Research*, vol. 57, no 12, pp. 3935–3953. DOI: 10.1080/00207543.2018.1443229.
24. Xu Y., Sun Y., Liu X., Zheng Y. (2019) A digital-twin-assisted fault diagnosis using deep transfer learning. *IEEE Access*, no 7, pp. 19990–19999. DOI: 10.1109/ACCESS.2018.2890566.
25. Borshchev A. (2013) *The Big Book of simulation modeling. Multimethod modeling with AnyLogic 6*. AnyLogic North America.

About the authors

Valery L. Makarov

Dr. Sci. (Phys.-Math.); Academician of Russian Academy of Sciences;
Academic Supervisor, Central Economics and Mathematics Institute, Russian Academy of Sciences, 47, Nakhimovsky Prospect, Moscow 117418, Russia;
E-mail: makarov@cemi.rssi.ru
ORCID: 0000-0002-2802-2100

Albert R. Bakhtizin

Dr. Sci. (Econ.); Corresponding Member of Russian Academy of Sciences;

Director, Central Economics and Mathematics Institute, Russian Academy of Sciences, 47, Nakhimovsky Prospect, Moscow 117418, Russia;

E-mail: albert@cemi.rssi.ru

ORCID: 0000-0002-9649-0168

Gayane L. Beklaryan

Cand. Sci. (Econ.);

Senior Researcher, Laboratory of Computer Modeling of Social and Economic Processes, Central Economics and Mathematics Institute, Russian Academy of Sciences, 47, Nakhimovsky Prospect, Moscow 117418, Russia;

E-mail: glbeklaryan@gmail.com

ORCID: 0000-0002-1286-0345

Andranik S. Akopov

Dr. Sci. (Tech.);

Professor, Department of Business Informatics, Graduate School of Business, National Research University Higher School of Economics, 20, Myasnitskaya Street, Moscow 101000, Russia;

Chief Researcher, Laboratory of Dynamic Models of Economy and Optimization, Central Economics and Mathematics Institute, Russian Academy of Sciences, 47, Nakhimovsky Prospect, Moscow 117418, Russia;

E-mail: aakopov@hse.ru

Оценка состояния объекта управления на основе универсального комплексного индикатора с использованием структурированных и неструктурированных данных

Т.К. Богданова 
E-mail: tanbog@hse.ru

Л.В. Жукова 
E-mail: lvzhukova@hse.ru

Национальный исследовательский университет «Высшая школа экономики»
Адрес: 101000, г. Москва, ул. Мясницкая, д. 20

Аннотация

Оценивание состояния объекта управления с помощью индикаторов, являющихся инструментами управления и контроля, активно используется во многих сферах экономики. Как правило, такие индикаторы строятся на основе внутренних данных. Однако с ростом объемов доступной открытой информации появляются алгоритмы оценки состояния определенных объектов управления с использованием открытых структурированных данных. Недостатком этих моделей является их узкая специализация и привязка только к структурированным, а иногда и строго официальным данным, которые, как правило, имеют редкую периодичность публикации. Это не позволяет отслеживать изменение состояния объекта в разные моменты времени. Авторами предложена концепция построения универсального комплексного индикатора (universal complex indicator, UCI) экспресс-оценки состояния объекта управления в различных видах деятельности: банковской, образовательной, производственной и т.д. Отличие предлагаемой концепции заключается в том, что в ней в качестве отправной точки выступают требования контролирующих органов, в то время как в большинстве российских и зарубежных исследований индикаторы строятся, прежде всего, в интересах инвесторов. Также предлагается использовать не только структурированные, но и неструктурированные данные, отслеживая динамику состояния объекта управления. Для определения значений UCI на основе различных эконометрических моделей и методов рассчитываются компоненты, характеризующие требования контролирующих органов к объекту управления. На основе этих требований с помощью таблицы истинности определяется значение UCI. Предложенная концепция апробирована для построения экспресс-оценки финансового состояния 108 банков за период с 1 января 2018 по 1 февраля 2020 года. В соответствии с требованиями ЦБ РФ были получены значения трех компонент UCI и рассчитано его значение для каждого банка. Прогностическая способность построенной модели, апробированной на трех банках тестовой выборки, была подтверждена согласованностью экспресс-оценки с их фактическим состоянием в марте 2020 года.

Ключевые слова: экспресс-оценка; универсальный комплексный индикатор; логическая функция; таблица истинности; объект управления; коммерческий банк; требования контролирующих органов; финансовое состояние; структурированные данные; неструктурированные данные; эконометрическая модель.

Цитирование: Богданова Т.К., Жукова Л.В. Оценка состояния объекта управления на основе универсального комплексного индикатора с использованием структурированных и неструктурированных данных // Бизнес-информатика. 2021. Т. 15. № 2. С. 21–33. DOI: 10.17323/2587-814X.2021.2.21.33

Введение

В современных экономических условиях роль грамотного управления в деятельности организации любого уровня нельзя недооценивать. При оптимальном управлении экономятся ресурсы, повышается эффективность деятельности, ускоряется возврат средств и снижаются риски финансовых потерь. В последние годы наблюдается пристальное внимание к способам и методам управления, а также к информации, на основе которой принимаются решения. Чем подробнее и качественнее используемая информация, тем точнее принятые на ее основе управленческие решения. Управление не существует без мониторинга и контроля состояния объекта управления [1]. О проблеме управления экономическими объектами в последние годы было написано достаточно много как российскими, так и зарубежными исследователями [2].

Проблемы совершенствования управления различными объектами были затронуты во многих научных публикациях последнего десятилетия. Например, в статье [3] автором подняты вопросы совершенствования инструментария планирования социально-экономического развития и оценки объектов управления в условиях активного формирования инфраструктуры информационного общества, а также дана оценка влияния цифровых технологий на содержание управленческих задач при изменчивости объектов и субъектов управления в крупномасштабных организационных системах. Отмечается недостаточно развитый уровень вовлеченности больших данных в управление на тактическом уровне [4]. Автор предлагает стандартизацию технологий сбора, передачи, хранения и обработки больших объемов данных, а также построение компьютерных математических моделей обеспечения многоуровневой системы государственного стратегического планирования [5, 6].

В работе [7] автор, анализируя исследования в области управления и контроля, отмечает, что в условиях ужесточения конкуренции эффективным инструментом выживания и приспособляемости

становится внутренний контроль. При этом отмечается слабо развитая в настоящее время система внутреннего контроля большинства предприятий. Автор делает вывод, что руководители большинства фирм считают внедрение внутреннего контроля затратным и трудоемким, при этом они недооценивают всей важности этой системы.

В работе [1] автор выделяет ряд значимых проблем управления, касающихся контрольных функций организации. Так, при проведении контроля акцент проверяющих органов делается, как правило, только на аудите финансово-хозяйственной деятельности. В результате контрольная деятельность превращается в перепроверку учетных операций, включая арифметический пересчет. В то же время оценка внешних неколичественных показателей, таких как уровень популярности, лояльность потребителей, качество объекта управления и т.п. остается вне внимания контрольных органов.

Анализ методов и моделей, предложенных российскими и зарубежными исследователями, преимущественно опирается на информацию, поступающую с определенной периодичностью, часто с существенным лагом относительно текущего момента. Вместе с тем контролирующим органам необходимо получать достоверную оценку состояния объекта управления на текущий момент времени, что позволит своевременно предпринять необходимые меры для предупреждения негативного развития событий.

Данная проблема может быть решена проведением экспресс-анализа текущего состояния объекта управления, опирающегося на использование как формализованной, так и неформализованной, как структурированной, так и неструктурированной информации. Результатом экспресс-анализа является получение экспресс-оценки состояния объекта управления. Экспресс-оценка состояния объекта управления — это предварительное исследование, во многом определяющее целесообразность проведения дальнейшего, более углубленного анализа.

Учитывая вышесказанное, цель настоящего исследования заключается в разработке универсального комплексного индикатора экспресс-оценки соответствия состояния экономического объекта управления заявленным требованиям со стороны регуляторов или соответствующих служб, на основе структурированных и неструктурированных данных, получаемых из сети интернет.

Основными задачами исследования являются:

- ♦ разработка концепции экспресс-оценки соответствия состояния объекта управления заявленным требованиям текущего контроллинга;
- ♦ разработка алгоритма построения универсального комплексного индикатора экспресс-оценки соответствия состояния объекта управления заявленным требованиям;
- ♦ апробация разработанной концепции на примере расчета универсального комплексного индикатора экспресс-оценки состояния российского коммерческого банка.

Актуальность исследования обусловлена следующими факторами:

- ♦ необходимостью разработки инструмента экономико-математического моделирования на основе открытых данных для текущего мониторинга, анализа и прогноза состояния объекта управления;
- ♦ потребностью в текущей оценочной модели, которая не зависит от момента публикации статистической отчетности;
- ♦ проведением экспресс-оценки состояния объекта управления (на примере банка) по открытым данным;
- ♦ необходимостью отслеживания динамики состояния объекта управления на основе структурированных и неструктурированных данных.

1. Постановка задачи исследования

Предлагаемый подход призван обеспечить решение проблемы трансформации существующей системы управления и контроля, а также развитие системы внутреннего контроля в условиях цифровизации экономики [7, 8]. Как правило, официальная статистическая отчетность публикуется в открытом доступе с определенной периодичностью и соответствующим лагом, что затрудняет оперативное реагирование в форс-мажорных ситуациях.

Подход к построению комплексного индикатора на основе использования открытых структурированных данных об объекте управления в текущий момент времени для оценки степени износа береговой зоны предложен в работе [9]. В работе сформулирована модель расчета комплексного индикатора на основе получаемой из публичных карт и материалов сообщений структурированной информации об использовании населением территории в качестве необустроенных диких пляжей, мини-отелей или мест общественного собрания.

Одна из проблем, поднимаемых авторами данной статьи, — это проблема постоянного контроля объектов управления с использованием официальной статистической отчетности, выходящей с большой периодичностью и с большим лагом. Например, в банковской сфере текущий контроль коммерческих банков со стороны Центрального банка затруднен из-за отсрочки поступления сведений о состоянии финансов банков. Это проблема рассмотрена в работе [10]. Авторы предлагают использовать открытые источники информации и на основе математической модели формировать комплексный индикатор с использованием таких данных [2, 11]. Изложенный в статье подход позволяет получать на основе открытых данных оперативную информацию о состоянии банка как объекта управления и ускорить принятие решений в его отношении. Результатом применения математической модели является вероятность отзыва лицензии (как одной из компонент модели оценки надежного банка) на основе публичной финансовой отчетности и учета волатильности внешней среды (изменений курсов валют, цен на нефть и т.д.) [12–14].

Отличительной особенностью данной работы является предложение использовать для оценки надежности банка не только структурированные, но и неструктурированные данные из открытых источников информации, такие как качественная и рейтинговая оценки общего состояния банка с точки зрения независимых российских рейтинговых агентств или тональность новостей с упоминанием банка.

Неструктурированные данные могут быть представлены в разных видах: текст, новостные сообщения, комментарии, фотографии [15]. С развитием интернет-технологий их объемы растут, и информация на основе таких данных может помочь в экспресс-оценке состояния объекта управления.

В отличие от рассмотренных выше работ, авторы предлагают концепцию построения универсального комплексного индикатора экспресс-оценки состояния объекта управления на основе открытых

структурированных и неструктурированных данных, полученных с официальных и неофициальных сайтов. Данные собираются как со статистических ресурсов, где они публикуются периодически, так и из интернет-ресурсов, регулярно обновляемых с разной периодичностью. В отличие от используемых в течение многих лет структурированных данных, неструктурированные данные представляют собой объединение разрозненной информации. Это позволяет конечным пользователям отслеживать на ранней стадии последствия принятых решений, что обеспечивает минимизацию рисков. Учет неструктурированных данных при оценке объектов управления дает возможность оперативного мониторинга текущего состояния объекта управления с учетом качественных характеристик (репутация, отзывы, тенденции) на основании данных, полученных бесконтактным методом из открытых источников.

Для обработки методами машинного обучения, математической статистики и математическими моделями неструктурированная информация предварительно структурируется в виде набора некоторых показателей методами математической статистики, машинного обучения или эконометрики [16–18].

Например, в докладе авторов на конференции [19] для экспресс-оценки состояния банка, отличного от стабильного, предложили использовать в построении комплексного индикатора одну из компонент, рассчитанную на основе неструктурированных данных (текстового содержания новостей), выделяя из них негативные упоминания названия банка как отдельного признака, влияющего на вероятность неустойчивости его состояния.

2. Концепция построения универсального комплексного индикатора экспресс-оценки состояния экономического объекта управления

Предлагаемая концепция формирования универсального комплексного индикатора (universal complex indicator, UCI) экспресс-оценки состояния объекта управления позволяет не только оценить текущее, но и спрогнозировать будущее состояние такого объекта управления [19–21].

Универсальный комплексный индикатор (UCI) может быть использован в следующих целях:

- ◆ для оценки близости состояния объекта управления к критическому уровню, установленному контролирующим органом;

- ◆ для анализа тенденции изменения состояния объекта управления на основе структурированных и неструктурированных данных;
- ◆ для сбора и анализа оперативной информации о состоянии объекта управления;
- ◆ для оценки достоверности отчетных данных о состоянии объекта управления;
- ◆ как инструмент, позволяющий сравнивать состояние разномасштабных объектов управления;
- ◆ для учета внешних по отношению к объекту управления характеристик его деятельности (наличие фактических мини-отелей на побережье, тональность новостей о директоре банка, комфортность инфраструктуры образовательного учреждения, активность использования социальных сетей респондентом и т.п.) [22–24].

Преимущество использования открытых данных — это возможность получать информацию с любой периодичностью (не привязываясь к регулярности обновления официально публикуемой статистической отчетности), расширять и проверять соответствие фактического состояния объекта управления официальным данным.

Концептуальная схема построения UCI экспресс-оценки состояния объекта управления для мониторинга и контроля с целью предотвращения развития негативных событий представлена на *рисунке 1*.

Построение универсального комплексного индикатора предусматривает следующие этапы.

Этап 1. На первом (начальном) этапе на основе информации о требованиях контролирующего органа к состоянию объекта управления определяются компоненты UCI. Такая информация может быть получена из постановлений, распоряжений, законов и других нормативных актов. В качестве контролирующего органа может выступать либо внешняя организация (Центральный банк, министерство), либо внутренний контрольный орган самой организации.

В результате реализации алгоритма формируется универсальный комплексный индикатор (UCI), значения которого могут быть целочисленными, представляющими собой бинарные или категориальные величины. Эти значения отражают состояние объекта управления относительно удовлетворительного, определяемого контролирующим органом. Так, примером бинарных значений универсального комплексного индикатора является: 0 — нет признаков незаконной деятельности организации, 1 — есть признаки незаконной деятельности организации.

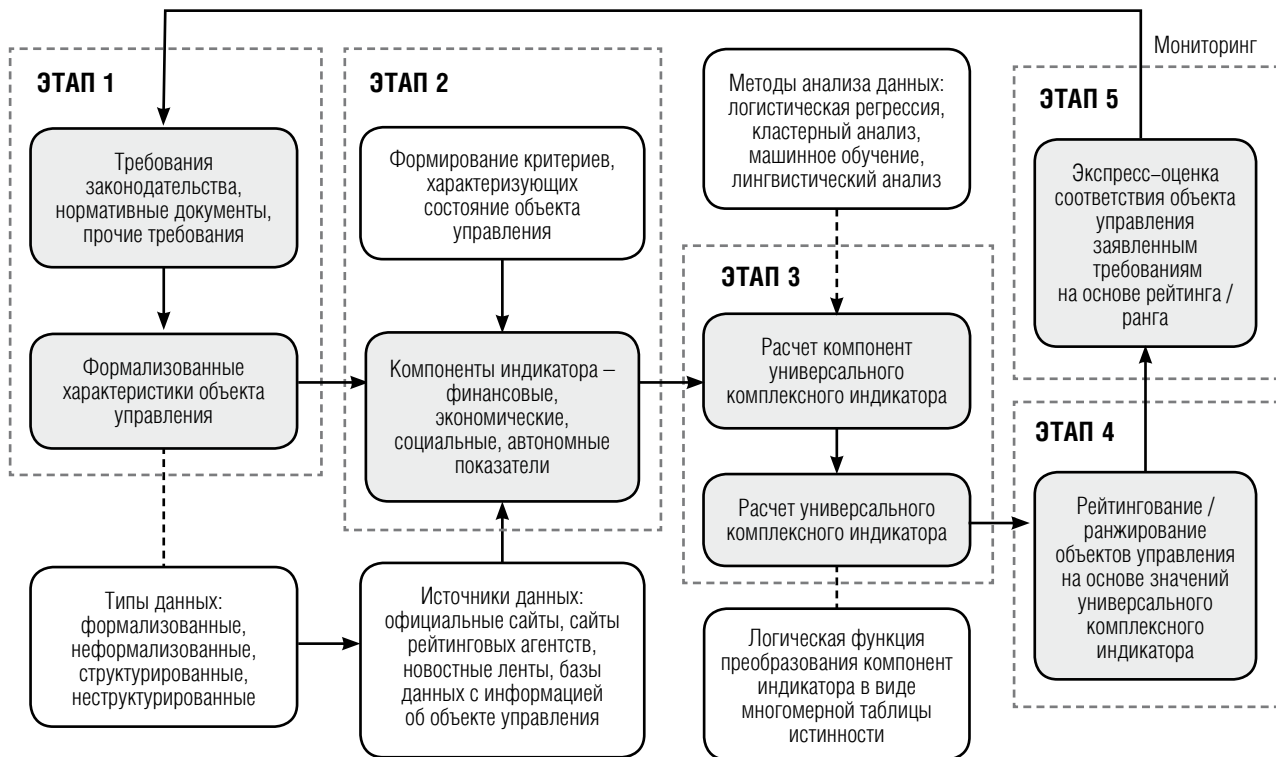


Рис. 1. Концептуальная схема построения универсального комплексного индикатора (UCI)

Чаще всего используется категориальная шкала с тремя категориями, например, 1 означает стабильное состояние, 2 – нестабильное состояние и 3 – состояние, требующее более пристального изучения.

Этап 2. После формирования критериев, характеризующих состояние объекта управления, которые в совокупности и должны описываться UCI, определяются компоненты индикатора, позволяющие оценить состояние различных характеристик объекта управления: экономическое состояние, вероятность наступления состояния объекта, отличного от стабильного, медийная активность, связанная с объектом, характеристики местоположения объекта, прогноз финансового или экономического состояния объекта и т.п. Затем формируется перечень компонент UCI и определяются данные из открытых источников, из которых может быть получена информация для математического моделирования и расчета компонент. Неструктурированные данные предварительно обрабатываются для возможности их использования в качестве расчетных показателей.

Этап 3. С помощью математических методов (логистическая регрессия, кластерный анализ, регрессионный анализ, статистический анализ) на основе

имеющихся данных рассчитываются компоненты UCI, представляющие собой численные значения, с помощью которых ранжируются все объекты управления.

На основе полученных значений компонент UCI определяется значение универсального комплексного индикатора экспресс-оценки соответствия состояния объекта управления заявленным требованиям. Для этого значения компонент в качестве аргументов подставляются в логическую функцию, агрегирующую различные значения компонент UCI.

Рассмотрим такую логическую функцию. Пусть от контролирующего органа поступило n требований, сформулированных в виде n критериев, и для универсального комплексного индикатора сформировано k компонент I_j , $j = 1, \dots, k$, принимающих значения из множества X (обычно $X \in \mathbf{R}$).

При этом одно требование к состоянию объекта управления от контролирующего органа может быть описано несколькими компонентами UCI, т.е. $k \geq n$. Универсальный комплексный индикатор (UCI) принимает значения из некоторого заданного дискретного множества Y , т.е. $y_i \in Y$, где $i = 1, \dots, p$.

Логическая функция $\Psi(I_1, I_2, \dots, I_k)$ задается та-

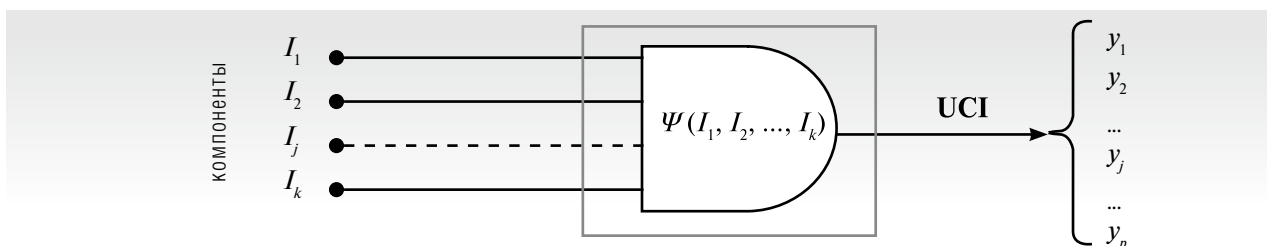


Рис. 2. Формирование UCI с использованием логической функции

блицей истинности (таблица 1), вводимой на основании важности критериев, полученных от контролирующего органа (рисунок 2).

Каждая компонента I_j может принимать Q_j значений. Значения компоненты $I_j^{q_j} \in X_j$, где $j = 1, \dots, k$; $q_j = 1, \dots, Q_j$; $X_j \in X$.

Общее количество комбинаций M всех возможных значений всех компонент k определяется следующим образом:

$$M = \prod_{j=1}^k Q_j. \quad (1)$$

В общем случае таблица представляет собой матрицу размерностью $M \times k$. Первая строка матрицы (сочетание s_1) представляется первыми значениями каждой из k компонент I_j^1 . Последняя строка матрицы представляется максимальными значениями каждой из k компонент $I_j^{(Q_j)}$. Строки, находящиеся между первой и последней строками, могут

формироваться различным образом, но так чтобы осуществлялся полный перебор всех возможных комбинаций значений, принимаемых каждой из компонент I_j .

Этап 4. На следующем этапе после вычисления значения UCI осуществляется ранжирование (рейтингование) объектов управления.

Этап 5. После этапа ранжирования производится оценка соответствия объекта управления заявленным требованиям в соответствии с полученными рангами (рейтингами). Исходя из полученных значений на текущий момент, вырабатываются рекомендации для надзорных органов относительно объектов управления. По мере появления новой информации в течение периода мониторинга группы объектов управления регулярно производится обновление данных из открытых источников, и на обновленных данных производится переоценка модели.

Таблица 1.

Общий вид многомерной таблицы истинности UCI

s_m – номер сочетания, $s_m = 1, \dots, M$	Компоненты универсального комплексного индикатора UCI					Значение универсального комплексного индикатора UCI
	I_1	...	I_j	...	I_k	
s_1	I_1^1	...	I_j^1	...	I_k^1	$y_{s_1} = \psi(I_1^1, \dots, I_j^1, \dots, I_k^1); y_{s_1} \in Y$
s_2	I_1^2	...	I_j^1	...	I_k^1	$y_{s_2} = \psi(I_1^2, \dots, I_j^1, \dots, I_k^1); y_{s_2} \in Y$
...	...	...	...	...	...	...
s_m	$I_1^{q_1}$	...	$I_j^{q_j}$	...	$I_k^{q_k}$	$y_{s_m} = \psi(I_1^{q_1}, \dots, I_j^{q_j}, \dots, I_k^{q_k}); y_{s_m} \in Y$
...	...	...	...	...	...	...
s_M	$I_1^{Q_1}$	...	$I_j^{Q_j}$	...	$I_k^{Q_k}$	$y_{s_M} = \psi(I_1^{Q_1}, \dots, I_j^{Q_j}, \dots, I_k^{Q_k}); y_{s_M} \in Y$

3. Расчет универсального комплексного индикатора экспресс-оценки состояния коммерческого банка

Предложенная концепция построения UCI экспресс-оценки соответствия состояния экономического объекта управления заявленным требованиям со стороны регуляторов была апробирована для экспресс-оценки состояния коммерческого банка. Контрольным органом в данном случае является ЦБ РФ — надзорный орган в банковской сфере.

Построение математической модели UCI экспресс-оценки состояния коммерческого банка в соответствии с рассмотренной выше концепцией включает пять этапов.

Этап 1. Формирование целевых показателей оценки состояния банка на основе требований к показателям ликвидности и надежности банка, полученных из нормативов и законодательных актов ЦБ РФ. В соответствии с установленными ЦБ РФ нормативами и требованиями к коммерческому банку были сформированы следующие характеристики:

- ♦ вероятность финансового банкротства коммерческого банка;
- ♦ финансовая стабильность коммерческого банка;
- ♦ качественная оценка деятельности коммерческого банка экспертами — участниками рынка.

Для оценки сформированных характеристик рассчитываются значения компонент UCI.

Этап 2. Определение компонент UCI экспресс-оценки состояния банка:

- ♦ **вероятность финансового банкротства коммерческого банка (характеристика ЦБ РФ).** Для определения этой характеристики рассчитывается компонента 1 — прогнозная вероятность наступления банкротства коммерческого банка, оцененная с помощью логистической регрессии, по финансовым и экономическим факторам;
- ♦ **финансовая стабильность коммерческого банка (характеристика ЦБ РФ).** Для определения этой характеристики рассчитывается компонента 2 — косвенный признак нестабильности коммерческого банка, полученный на основе оценки принадлежности банка к группе «неблагополучных». Однородные группы банков по признакам стабильности выявлены методом кластеризации [25];
- ♦ **качественная оценка деятельности коммерческого банка экспертами — участниками рын-**

ка (характеристика ЦБ РФ). Для определения этой характеристики рассчитывается компонента 3 — наличие негативной оценки общего состояния банка со стороны участника рынка — экспертного агентства «Эксперт РА». Индикатор построен с помощью качественной оценки на основе семантического анализа текста новостей с упоминанием банка.

Этап 3. Построение UCI экспресс-оценки состояния коммерческого банка. Для этого была использована обучающая выборка из 108 действующих коммерческих банков. Информация была получена из восьми внешних источников, включая сайт ЦБ РФ, рейтинговые агентства, справочники и информационно-аналитические порталы. Требования к состоянию банка были взяты из нормативов регулятора — ЦБ РФ.

Для характеристики стабильности финансового состояния коммерческого банка и его рыночного поведения оценивалась существенность колебаний в динамике изменения банковских ставок без видимых причин. Основные показатели, характеризующие ставки банков по вкладам, были взяты за период с 01.01.2018 по 01.02.2020 гг., показатели рейтинговых агентств — с 2017 по 2020 годы, а общие характеристики коммерческих банков по данным сайта Банка России и других открытых источников — по состоянию на февраль 2020 года.

В модели расчета UCI было использовано более 60 исходных показателей следующего вида:

- ♦ ставки каждого коммерческого банка по депозитным вкладам на разные сроки;
- ♦ показатели, характеризующие состояние банка: его активы и значения нормативов ($H_1 - H_7$);
- ♦ макроэкономические показатели внешней среды: ключевая ставка, средние ставки по всем банкам РФ и по 30 крупнейшим банкам РФ;
- ♦ показатели рейтинговых агентств за период с 2017 по 2020 годы;
- ♦ семантическая (текстовая) информация об упоминании каждого банка в новостной ленте рейтингового агентства.

На основе этих исходных данных были рассчитаны показатели центральной тенденции и разброса:

- ♦ медиана и дисперсия отклонений ставок по депозитным вкладам для каждого коммерческого банка относительно макроэкономических показателей: ключевой ставки, среднебанковской ставки по всем банкам и по 30 крупнейшим банкам России;

- ♦ относительные показатели динамики ставок, их медианы и дисперсии;
- ♦ относительное и абсолютное место банка в рейтинге банков;
- ♦ негативное упоминание банка в новостной ленте на сайте рейтингового агентства.

Критерием проверки качества модели служил статус коммерческого банка на момент сбора информации, то есть являлся ли он действующим или ликвидированным.

Для получения значений UCI компоненты I_1 , I_2 и I_3 были рассчитаны следующим образом.

Компонента 1 (I_1) – прогнозная вероятность наступления банкротства коммерческого банка. Для получения оценки вероятности отзыва лицензии у коммерческого банка построена логистическая регрессионная модель на основе исходных данных и расчетных показателей. К числу таких показателей относятся: место банка в рейтинге, волатильность банковских ставок, качественный расчетный показатель, характеризующий восприятие экспертным сообществом деятельности коммерческого банка, на основе обработки неструктурированных текстовых данных – упоминание о понижении рейтинга коммерческого банка в новостной ленте на сайте рейтингового агентства. Зависимая переменная Y – бинарная переменная, принимающая два значения: 0, если текущий статус банка – «действующий», и 1 – если он ликвидирован.

Оцененная логистическая регрессионная модель имеет вид:

$$P(Y = 1) = \frac{1}{1 + e^{-z}}, \quad (2)$$

$$z = -16,7 + 1,4955 \cdot x_1 + 0,024 \cdot x_2 + 1,128 \cdot x_3 + 0,389 \cdot x_4, \quad (3)$$

где x_1 – понижение в рейтинге;

x_2 – место в рейтинге;

x_3 – медиана ставок по депозитам;

x_4 – медиана отклонений ставок по депозитам до востребования от среднерыночной ставки.

Все коэффициенты модели значимы на 5% уровне, показатели чувствительности (т.е. доля ликвидированных банков) и специфичности (доля действующих банков) модели высокие, соответственно 85,9% и 73,3%.

По результатам работы модели с помощью разбиения всех полученных значений на три интервала была сформирована компонента 1 UCI, характеризующая вероятность отзыва лицензии в течение полугодия после момента оценки:

♦ низкая: $P(Y = 1) \leq 0,25$;

♦ средняя: $0,25 < P(Y = 1) < 0,35$;

♦ высокая: $P(Y = 1) \geq 0,35$.

Границы отсечения выбирались на основе оптимизации чувствительности и специфичности следующим образом. Верхняя граница, равная 0,35, была получена при одновременном достижении максимальных значений чувствительности и специфичности модели (85,9% и 73,3%). Нижняя граница, равная 0,25, соответствует значению специфичности, равному 80%.

Компонента 2 (I_2) – Оценка принадлежности банка к кластеру «неблагополучных». При кластеризации обучающей выборки методом BIRCH (двухшаговый метод) по основным финансовым показателям коммерческого банка по данным о значениях семи основных нормативов банков ($H_1 - H_7$), по месту коммерческого банка в рейтинге банков рейтингового агентства и изменению этого рейтинга в ретроспективе было получено пять однородных кластеров.

Наиболее интересным оказался первый кластер, составляющий 14% от обучающей выборки. В этом кластере присутствовало 30% ликвидированных банков. У всех банков этого кластера место в рейтинге банков находилось ниже 200. Также было отмечено снижение медианы нормативов каждого коммерческого банка за ретроспективный период и снижение медианы дисперсии ставок по депозитным вкладам за предыдущий период.

По результатам работы модели была сформирована компонента 2 UCI – принадлежность к первому кластеру, как признак «неблагополучия» коммерческого банка.

Компонента 3 (I_3) – наличие негативной оценки деятельности банка со стороны участника рынка – экспертного агентства. Эта компонента формируется как бинарный показатель принадлежности к одному из двух множеств. Компонента 3 принимает значение 1, если есть хотя бы одно негативное упоминание наименования коммерческого банка в новостной ленте на сайте рейтингового агентства за текущий год, и значение 0, если нет ни одного негативного упоминания.

Из компонент UCI с помощью таблицы истинности (таблица 1) были рассчитаны значения универсального комплексного индикатора (UCI) для обучающей и тестовой выборок, приведенные в таблице 2.

Таблица 2.

Значения UCI оценки состояния банка

Компоненты универсального комплексного индикатора (UCI)			Значение (y_j) универсального комплексного индикатора (UCI)
Компонента 1 (I_1): Вероятность отзыва лицензии	Компонента 2 (I_2): Принадлежность к кластеру	Компонента 3 (I_3): Наличие негативных упоминаний на сайте рейтингового агентства	
Высокая (более 0,35)	Принадлежность к кластеру не имеет значения	Наличие или отсутствие негативных упоминаний не имеет значения	3: Требуется пристальное внимание к банку
Средняя (0,25–0,35)	Принадлежит к кластеру 1	1 (наличие негативных упоминаний)	
Средняя (0,25–0,35)	Принадлежит к кластеру 1	0 (отсутствие негативных упоминаний)	2: Требуется внимание к банку
Средняя (0,25–0,35)	Не принадлежит к кластеру 1	1 (наличие негативных упоминаний)	
Низкая (менее 0,25)	Принадлежность к кластеру не имеет значения	1 (наличие негативных упоминаний)	
Низкая (менее 0,25)	Принадлежность к кластеру не имеет значения	0 (отсутствие негативных упоминаний)	1: Внимание к банку не требуется
Средняя (0,25–0,35)	Не принадлежит к кластеру 1	0 (отсутствие негативных упоминаний)	

Таблица 3.

Значение UCI для банков тестовой выборки

Наименование банка	Компонента 1 (I_1): Вероятность отзыва лицензии	Компонента 2 (I_2): Принадлежность к кластеру	Компонента 3 (I_3): Наличие негативных упоминаний на сайте рейтингового агентства	Значение (y_j) универсального комплексного индикатора (UCI)
ПАО КБ «ПФС–Банк»	Высокая (0,844)	Не принадлежит к кластеру 1	0	3: Требуется пристальное внимание к банку
АКБ «Пересвет» (ПАО)	Низкая (0,01)	Принадлежит к кластеру 1	1	2: Требуется внимание к банку
АО КБ «Ситибанк»	Низкая (0,007)	Не принадлежит к кластеру 1	0	1: Внимание к банку не требуется

Таблица 4.

Сопоставление расчетного значения UCI и текущего статуса коммерческого банка

Наименование банка	Значение (y_j) универсального комплексного индикатора (UCI)	Текущий статус коммерческого банка
ПАО КБ «ПФС–Банк»	3, требуется пристальное внимание к банку	Лицензия отозвана в марте 2020 г.
АКБ «Пересвет» (ПАО)	2, требуется внимание к банку	Действующий. Наблюдается неоднократное нарушение обязательных нормативов регулятора в процессе санации банка
АО КБ «Ситибанк»	1, не требуется внимание к банку	Действующий

Этап 4. Ранжирование коммерческих банков учающей выборки на основе рассчитанных значений UCI. Результатом выполнения этого этапа является список коммерческих банков, отсортированный по убыванию значений UCI.

Этап 5. Аprobация разработанной модели расчета UCI была проведена на тестовой выборке из трех коммерческих банков. Результаты моделирования и рекомендации для коммерческих банков из тестовой выборки приведены в таблице 3.

Как видно из таблицы 4, результаты апробации разработанной модели на тестовой выборке из трех коммерческих банков подтверждаются текущим статусом коммерческих банков по состоянию на март 2020.

Таким образом, предложенная на основе анализа открытых структурированных и неструктурированных данных концепция экспресс-оценки состояния коммерческого банка позволяет оценить вероятность наступления неудовлетворительного состояния банка. Полученные результаты позволяют принять превентивные меры поддержки или защиты интересов вкладчиков со стороны надзорного органа (ЦБ РФ).

Заключение

В статье предложены концепция и алгоритм построения универсального комплексного индикато-

ра, позволяющие получить экспресс-оценку состояния объекта управления со стороны контрольных органов, с целью предотвращения развития негативных событий.

Новизной предложенного подхода является возможность применения универсального комплексного индикатора (universal complex indicator, UCI) для экспресс-оценки объектов управления в различных предметных областях и сферах деятельности, с использованием в качестве исходной информации структурированных и неструктурированных данных из открытых источников.

Сформирована информационная база исследования для экспресс-оценки состояния российских коммерческих банков, включающая 111 банков и 79 финансовых и нефинансовых показателей за период с 01.01.2018 по 01.02.2020 гг.

Проведена апробация разработанной концепции построения универсального комплексного индикатора (UCI) экспресс-оценки состояния трех российских коммерческих банков и выявлен банк с состоянием, не соответствующим предъявляемым со стороны ЦБ РФ требованиям. Выводы о статусе трех коммерческих банков по состоянию на март 2020 из тестовой выборки, полученные на основе анализа значений UCI, в дальнейшем были подтверждены информацией, размещенной на сайте ЦБ РФ. ■

Литература

1. Муллахметов Х.Ш. Проблемы организации управленческого контроля // Экономическая наука современной России. 2007. № 4 (39). С. 59–69.
2. Полшков Ю.Н. Прикладные эконометрические методы анализа рисков при управлении хозяйственным комплексом региона и уровнем жизни его населения // Вестник Института экономических исследований. 2019. №1 (13). С. 12–18.
3. Писарева О.М. Анализ состояния и характеристика потенциала развития инструментария стратегического планирования в условиях цифровой трансформации экономики и управления // МИР (Модернизация. Инновации. Развитие). 2018. Т. 9. № 4. С. 502–529. DOI: 10.18184/2079-4665.2018.9.4.502-529.
4. Mayer-Sch nberger V., Cukier K. Big data: A revolution that will transform how we live, work, and think. Houghton Mifflin Harcourt, 2013.
5. Big data: The next frontier for innovation, competition, and productivity / J. Manyika [et al.]. McKinsey Global Institute, 2011.
6. Горев А.И., Горева Е.Г. О применимости существующих алгоритмов обработки данных к Big Data // Математические структуры и моделирование. 2020. № 1 (53). С. 139–143. DOI: 10.24147/2222-8772.2020.1.139-143.
7. Громова С.В. Проблемы организации внутреннего контроля на предприятии // Молодой ученый. 2016. № 18 (122). С. 240–242.
8. Гайдук Е.А. Применение балансового и оптимизационного моделирования при принятии управленческих решений на муниципальном уровне. Диссертация на соискание ученой степени кандидата экономических наук. Новосибирск, 2012.
9. Кирюшина А.А., Жукова Л.В., Чикина Л.Г. Использование больших данных в оценке степени загрязнения прибрежной зоны морского побережья курортных регионов // Тезисы докладов Всероссийской научной конференции «Моря России: исследования береговой и шельфовой зон». Севастополь, 21–25 сентября 2020. С. 409–410.

10. Биджоян Д.С., Богданова Т.К. Концепция моделирования и прогнозирования вероятности отзыва лицензии российских банков // Экономическая наука современной России. 2017. Т. 79. № 4. С. 88–103.
11. Марков С.М., Маркова А.С. Целевые показатели (индикаторы) как критерии комплексной оценки эффективности реализации государственных программ поддержки малого и среднего предпринимательства // Вестник Псковского государственного университета. Серия: Экономика. Право. Управление. 2017. № 5. С. 132–137.
12. Головань С.В., Евдокимов М.А., Карминский А.М., Пересецкий А.А. Модели вероятности дефолта российских банков II. Влияние макроэкономических факторов на устойчивость банков / Препринт #WP/2004/043. М.: Российская Экономическая Школа, 2004.
13. Пересецкий А.А. Модели причин отзыва лицензии российских банков / Препринт #WP/2010/085. М.: Российская Экономическая Школа, 2010.
14. Малых Н.И., Проданова Н.А. Современные подходы к оценке стоимости коммерческого банка // Статистика и экономика. 2016. №6. С. 79–84. DOI: 10.21686/2500-3925-2016-6-79-84.
15. Бородин О.Н. Извлечение информации из полнотекстовых источников данных // Агроинженерия. 2008. № 1. С. 42–43.
16. Zhang T., Ramakrishnan R., Livny M. (1996) BIRCH: An efficient data clustering method for very large databases // Proceedings of the 1996 ACM SIGMOD International Conference on Management of Data (SIGMOD '96), Montreal, Canada, 4–6 June 1996, P. 103–114. DOI: 10.1145/233269.233324.
17. Айвазян С.А., Афанасьев М.Ю. Оценка мероприятий, направленных на управление факторами неэффективности производства // Прикладная эконометрика. 2007. № 4(8). С. 27–41.
18. Пересецкий А.А. Эконометрические методы в дистанционном анализе деятельности российских банков. М.: ВШЭ, 2012.
19. Богданова Т.К., Жукова Л.В. Построение комплексного индикатора для оценки состояния российского коммерческого банка на основе структурированных и неструктурированных данных // Тезисы докладов конференции «43-е заседание международной научной школы-семинара «Системное моделирование социально-экономических процессов», Москва, 13–18 октября 2020 г.
20. Bidzhoyan D., Bogdanova T. Russian banks credit risk stress-testing based on the publicly available data // Advances in Intelligent Systems and Computing. 2019. Vol. 850. P. 262–271. DOI: 10.1007/978-3-030-02351-5_31.
21. Соболевская Ю.В. Контрольные органы: уровни компетенции. Вопросы государственного и муниципального управления. 2013. № 4. С. 113–130.
22. Карминский А.М., Костров А.В. Моделирование вероятности дефолта российских банков: расширенные возможности // Журнал Новой экономической ассоциации. 2013. № 1 (17). С. 64–86.
23. Поляков К.Л., Жукова Л.В. Опыт моделирования вероятности кредитного дефолта клиентов микрофинансовых организаций (на примере одной МФО) // Экономический журнал Высшей школы экономики. 2019. Т. 23. № 4. С. 497–523. DOI: 10.17323/1813-8691-2019-23-4-497-523.
24. Жукова Л.В., Кирюшина А.А., Ковальчук И.М., Рузаева А.В. Повышение результативности системы дистанционного образования с помощью машинного обучения и технологии блокчейн // Прикаспийский журнал: управление и высокие технологии. 2018. № 1 (41). С. 56–68.
25. Zhukova L., Polyakov K.L. Comparative analysis of predictive analytics models in classification problems // VI International Conference on Actual Problems of Systems and Software Engineering (APSSE 2019), Moscow, Russia, 12–14 November 2019, P. 162–169.

Об авторах

Богданова Татьяна Кирилловна

кандидат экономических наук, доцент;

доцент департамента бизнес-информатики, Высшая школа бизнеса, Национальный исследовательский университет «Высшая школа экономики», 101000, г. Москва, ул. Мясницкая, д. 20;

E-mail: tanbog@hse.ru

ORCID: 0000-0002-0018-2946

Жукова Людмила Вячеславовна

старший преподаватель департамента прикладной экономики, факультет экономических наук, Национальный исследовательский университет «Высшая школа экономики», 101000, г. Москва, ул. Мясницкая, д. 20;

E-mail: lvzhukova@hse.ru

ORCID: 0000-0003-1647-5337

Valuating the position of the control object based on a universal complex indicator using structured and unstructured data

Tatiana K. Bogdanova

E-mail: tanbog@hse.ru

Liudmila V. Zhukova

E-mail: lvzhukova@hse.ru

National Research University Higher School of Economics

Address: 20, Myasnitskaya Street, Moscow 101000, Russia

Abstract

Valuating the position of a controlled object using indicators which are management and control tools is widely used in many areas of the economy. Usually such indicators are based on internal data, however, as the volume of available open information grows, algorithms for valuation of the position of certain control objects and on open structured data are appearing. The disadvantage of these models is their narrow specialization and binding only to structured, and sometimes strictly official data, which, as a rule, have a rare publication frequency. This does not allow you to track the change in the position of the object at different times. The authors have proposed a concept for constructing a universal complex indicator (UCI) for express valuation of the position of a controlled object in various types of activity: banking, educational, industrial, etc. Another difference in the construction of the UCI is that the concept presented in the article assumes, as a reference point, to take into account the requirements of regulatory authorities, while in most Russian and foreign studies, indicators are built for the needs of investors. It is also proposed to use, along with structured and unstructured data, tracking the dynamics of changes in the position of the control object. To determine the UCI values on the basis of various econometric models and methods, the components that characterize the requirements of the control bodies to the control object are calculated; using them the UCI value is determined from the truth table. The concept proposed was tested to build an express valuation of the financial position of 108 banks for the period from 1 January 2018 to 1 February 2020. In accordance with the requirements of the Central Bank of the Russian Federation, the values of the three UCI components were obtained, and the value was calculated for each bank. The predictive ability of the constructed model, tested on three banks of the test sample, was confirmed by the consistency of the express valuation with their actual position in March 2020.

Key words: express valuation; universal complex indicator; logical function; truth table; control object; commercial bank; requirements of regulatory authorities; financial position; structured data; unstructured data; econometric model.

Citation: Bogdanova T.K., Zhukova L.V. (2021) Valuating the position of the control object based on a universal complex indicator using structured and unstructured data. *Business Informatics*, vol. 15, no 2, pp. 21–33. DOI: 10.17323/2587-814X.2021.2.21.33

References

1. Mullakhmetov Kh.Sh. (2007) Problems of management control organization. *Economics of Contemporary Russia*, no 4, pp. 59–69 (in Russian).
2. Polshkov Y.N. (2019) Applied econometric approaches to risk analysis in managing the region's economic complex and the living standard of its population. *Vestnik of Institute of Economic Research*, no 1, pp. 12–18 (in Russian).
3. Pisareva O.M. (2018) Analysis of the state and characteristics of the development potential of strategic planning tools in the digital transformation conditions of the economy and Management. *MIR (Modernization. Innovation. Research)*, vol. 9, no 4, pp. 502–529 (in Russian). DOI: 10.18184/2079-4665.2018.9.4.502-529.
4. Mayer-Schönberger V., Cukier K. (2013) *Big data: A revolution that will transform how we live, work, and think*. Houghton Mifflin Harcourt.

5. Manyika J., Chui M., Brown B., Bughin J., Dobbs R., Roxburgh C., Byers A.H. (2011) *Big data: The next frontier for innovation, competition, and productivity*. McKinsey Global Institute.
6. Gorev A.I., Goreva E.G. (2020) About the applicability of existing data processing algorithms to big data. *Mathematical Structures and Modeling*, no 1, pp. 139–143 (in Russian). DOI: 10.24147/2222-8772.2020.1.139-143.
7. Gromova S.V. (2016) Problems of organization of internal control at the enterprise. *Young Scientist*, no 8, pp. 240–242 (in Russian).
8. Gayduk E.A. (2012) *Application of balance and optimization modeling in making management decisions at the municipal level*. Doctoral thesis. Novosibirsk (in Russian).
9. Kiryushina A.A., Zhukova L.V., Chikina L.G. (2020) Using big data in assessing the degree of pollution of the coastal zone of the seacoast of resort regions. Proceedings of the *All-Russian Scientific Conference "Russian Seas: research of coastal and shelf zones"*, Sevastopol, 21–25 September 2020, pp. 409–410 (in Russian).
10. Bidzhoyan D.S., Bogdanova T.K. (2017) The concept of modeling and forecasting the probability of revoking a license of Russian banks. *Economics of Contemporary Russia*, vol. 79, no 4, pp. 88–103 (in Russian).
11. Markov S.M., Markova A.S. (2017) Key performance indicators as criteria for a comprehensive assessment of the effectiveness of the implementation of state programs to support small and medium-sized businesses. *Bulletin of Pskov State University. Series: Economics. Law. Management*, no 5, pp. 132–137 (in Russian).
12. Golovan S.V., Evdokimov M.A., Karminsky A.M., Peresetsky A.A. (2004) *Probability of default models of Russian banks II. Models and macroeconomic environment*. Working Paper #WP/2004/043. Moscow: New Economic School (in Russian).
13. Peresetsky A.A. (2010) Модели причин отзыва лицензии российских банков. *Modelling reasons for Russian bank license withdrawal*. Working Paper #WP/2010/085. Moscow: New Economic School (in Russian).
14. Malykh N.I., Prodanova N.A. (2016) Modern approaches to the assessment of the cost of commercial bank. *Statistics and Economics*, no 6, pp. 79–84 (in Russian). DOI: 10.21686/2500-3925-2016-6-79-84.
15. Borodin O.N. (2008) Extracting information from full-text data sources. *Agricultural Engineering*, no 1, pp. 42–43 (in Russian).
16. Zhang T., Ramakrishnan R., Livny M. (1996) BIRCH: An efficient data clustering method for very large databases. Proceedings of the *1996 ACM SIGMOD International Conference on Management of Data (SIGMOD '96)*, Montreal, Canada, 4–6 June 1996, pp. 103–114. DOI: 10.1145/233269.233324.
17. Aivazyan S.A., Afanasiev M.Yu. (2007) Evaluation of measures aimed at managing production inefficiency factors. *Applied Econometrics*, no 4, pp. 27–41 (in Russian).
18. Peresetsky A.A. (2012) *Econometric approach to off-site analysis of Russian banks*. Moscow: HSE (in Russian).
19. Bogdanova T.K., Zhukova L.V. (2020) Building a comprehensive indicator for assessing the state of a Russian commercial bank based on structured and unstructured data. Proceedings of the *43rd Meeting of the International Scientific School-Seminar on System Modeling of Socio-Economic Processes*, Moscow, 13–18 October 2020 (in Russian).
20. Bidzhoyan D., Bogdanova T. (2019) Russian banks credit risk stress-testing based on the publicly available data. *Advances in Intelligent Systems and Computing*, vol. 850, pp. 262–271. DOI: 10.1007/978-3-030-02351-5_31.
21. Sobolevskaya Yu.V. (2013) Control bodies: Competence levels. *Public Administration Issues*, no 4, pp. 113–130 (in Russian).
22. Karminsky A.M., Kostrov A.V. (2013) Modelling the default probabilities of Russian banks: Extended abilities. *Journal of the New Economic Association*, no 1, pp. 64–86 (in Russian).
23. Polyakov K.L., Zhukova L.V. (2019) Modeling the probability of credit default of clients of microfinance organizations: The case of one MFI. *HSE Economic Journal*, vol. 23, no 4, pp. 497–523. DOI: 10.17323/1813-8691-2019-23-4-497-523 (in Russian).
24. Zhukova L.V., Kiryushina A.A., Kovalchuk I.M., Ruzaeva A.V. (2018) Online education system efficiency improvement by the means of machine learning and blockchain technology. *Caspian Journal: Control and High Technologies*, no 1, pp. 56–68 (in Russian).
25. Zhukova L., Polyakov K.L. (2019) Comparative analysis of predictive analytics models in classification problems. Proceedings of the *VI International Conference on Actual Problems of Systems and Software Engineering (APSSE 2019)*, Moscow, Russia, 12–14 November 2019, pp. 162–169.

About the authors

Tatiana K. Bogdanova

Cand. Sci. (Econ.);

Associate Professor, Department of Business Informatics, Graduate School of Business, National Research University Higher School of Economics, 20, Myasnitskaya Street, Moscow 101000, Russia;

E-mail: tanbog@hse.ru

ORCID: 0000-0002-0018-2946

Liudmila V. Zhukova

Senior Lecturer, Department of Applied Economics, Faculty of Economic Sciences, National Research University Higher School of Economics, 20, Myasnitskaya Street, Moscow 101000, Russia;

E-mail: lvzhukova@hse.ru

ORCID: 0000-0003-1647-5337

Улучшение клиентского опыта взаимодействия с искусственным интеллектом путем соблюдения этических принципов

О.И. Долганова 
E-mail: oidolganova@fa.ru

Финансовый университет при Правительстве Российской Федерации
Адрес: 105187, г. Москва, ул. Щербаковская, д. 38

Аннотация

Интенсивное развитие и применение технологий искусственного интеллекта при организации взаимодействия с клиентами сопровождается такими сложностями, как нежелание клиента общаться с роботом, недоверие, боязнь, негативный клиентский опыт. Такие проблемы могут быть решены за счет соблюдения этических принципов использования искусственного интеллекта. В научных и практических работах по данной теме много рекомендаций общего характера, которые сложно применить на практике, или же наоборот, описываются методы для решения узкоспециализированной технической или управленческой задачи. Целью данной статьи является определение этических принципов и методов, соблюдение и реализация которых позволит повысить доверие клиента конкретной организации к системам искусственного интеллекта. В результате анализа и синтеза результатов научных и практических исследований, а также эмпирического опыта российских и зарубежных компаний выявлены основные области применения технологий искусственного интеллекта, влияющие на клиентский опыт. Сформулированы и систематизированы этические принципы, рекомендуемые к соблюдению бизнесом. Также определены основные методы, позволяющие реализовывать данные принципы на практике, снижая негативные эффекты от взаимодействия клиентов с искусственным интеллектом и повышая их доверие к компании.

Ключевые слова: этика; искусственный интеллект; этика искусственного интеллекта; клиентский опыт; этические принципы; машинное обучение; доверие; робот.

Цитирование: Долганова О.И. Улучшение клиентского опыта взаимодействия с искусственным интеллектом путем соблюдения этических принципов // Бизнес-информатика. 2021. Т. 15. № 2. С. 34–46.
DOI: [10.17323/2587-814X.2021.2.34.46](https://doi.org/10.17323/2587-814X.2021.2.34.46)

Введение

Экономический потенциал применения технологий искусственного интеллекта (ИИ) при взаимодействии с клиентами является значительным, однако есть и проблемы безопасности данных, прозрачности алгоритмов поведения машин и доверия к подобным инструментам со стороны клиентов. Именно поэтому вопросы цифровой этики все чаще поднимаются как в научной литературе, так и на практике. Проведенный анализ публикаций по исследуемой теме в Web of Science и Scopus показал, что в общем числе статей по запросу «AI Ethics» за период с 2001 по 2020 год более 57 % (483 работы) в Web of Science и 61% (587 работ) в Scopus приходится на публикации 2019 и 2020 годов. До 2019 года в среднем публиковалось около 30 работ в год, при этом более 40% публикаций принадлежит авторам из США и Великобритании. Также в число лидеров входят авторы из Австралии, Италии, Нидерландов и Канады. В сумме по двум базам по данной теме авторами из США опубликовано 528 работ, а российскими учеными – всего 14. Также важно отметить, что только чуть больше половины всех опубликованных работ относится к области компьютерных и социальных наук, бизнеса и экономики.

Такая тенденция наблюдается и в практико-ориентированных изданиях, а также материалах государственных и международных комитетов и экспертных советов [1–5]. Здесь предлагаются высокоуровневые подходы, принципы и методы решения данных проблем, которые сложно применить на практике на уровне компании. При этом больше половины руководителей компаний, использующих технологии ИИ, подчеркивают важность проблемы обеспечения их этичности и прозрачности [6]. Gartner также отмечает, что в ближайшие годы вопросы цифровой этики, как элемента корпоративной архитектуры, будут находиться на пике популярности [7].

В процессах маркетинга, продаж и послепродажного обслуживания искусственный интеллект может быть использован для решения совершенно разных задач. Примерами могут служить улучшение распознавания речи и анализ эмоционального состояния клиента, маршрутизация звонков, обработка обращений, обеспечение индивидуального подхода к каждому покупателю, поиск новых клиентов.

Термин «искусственный интеллект» не имеет единого устоявшегося определения. Многие исследователи считают, что понятие ИИ относится к программам, алгоритмам и системам, демонстрирующим

интеллект [8]. Однако такая формулировка вызывает много дискуссий на предмет того, как определить, что машина демонстрирует именно интеллект и какие ее действия доказывают ее разумность. Поэтому в данной работе будем придерживаться несколько иной точки зрения, подразумевающей, что информационные системы с ИИ построены на базе технологий машинного обучения и в них могут применяться инструменты роботизированной автоматизации процессов, обработки естественного языка, нейронные сети и методы глубокого обучения [9–11]. Такие программные решения позволяют интерпретировать имеющиеся данные, извлекать уроки, обучаться на них и адаптироваться под текущие потребности пользователя [12].

Применение подобных передовых цифровых технологий сопряжено с проблемами этического характера [11, 13, 14]. Клиенты не склонны доверять искусственному интеллекту. С точки зрения соблюдения этических норм они ожидают от подобных информационных систем больше, чем регламентировано текущими правовыми нормами [15]. Также для многих людей важно, чтобы взаимодействие с ними осуществлялось честно и прозрачно, только тогда они начинают доверять продавцу [16]. Именно доверие выступает в роли самого сильного фактора, влияющего на лояльность клиента [17–19]. Оно возникает в результате последовательного поведения компании, демонстрирующего ее честность и надежность. Важность этих аспектов в выстраивании взаимоотношений между продавцом и потребителем возрастает с каждым годом. Эту тенденцию, в частности, подтверждают результаты исследований КПМГ 2019–2020 годов [17, 20].

Вопросы улучшения клиентского опыта активно рассматриваются исследователями с точки зрения маркетинга, автоматизации продаж и послепродажного обслуживания, в том числе за счет использования технологий искусственного интеллекта [21, 22]. Однако их применение провоцирует возникновение этических проблем. Поэтому появилась острая потребность в исследовании данных вопросов и определении возможных путей снижения негативных последствий.

Целью данной работы является определение этических принципов и подходов к улучшению клиентского опыта взаимодействия с ИИ на этапе продаж и послепродажного обслуживания. В результате исследования планируется ответить на следующие вопросы:

- ♦ Какие принципы этики необходимо соблюдать для улучшения клиентского опыта взаимодействия с искусственным интеллектом?
- ♦ Какие шаги позволят снизить негативное отношение покупателей и потребителей к применению компанией технологий ИИ в процессах управления и реализации взаимоотношений с клиентами?

В рамках исследования был проведен анализ научных и практических публикаций, фреймворков, «белых книг» и аналитических отчетов, касающихся поставленных вопросов с точки зрения маркетинга, бизнеса, психологии, этики и информационных технологий.

Опираясь на классификацию систем искусственного интеллекта, используемых в клиентском обслуживании [23], можно выделить два сценария их применения: 1) робот как помощник человека, который обслуживает клиента и 2) робот как замена человека, обслуживающего клиента. В данной работе в основном рассматривается второй вариант. Это позволит сузить область исследования и сконцентрироваться на проблемах, возникающих именно при таком типе взаимодействия с клиентом.

1. Область применения технологий искусственного интеллекта в реализации взаимодействия с клиентами

В сфере управления и реализации взаимодействий с клиентами технологии искусственного интеллекта могут быть использованы для решения таких задач, как автоматизация процессов продаж, обработка обращений и жалоб, поиск и привлечение новых клиентов, повышение лояльности и удержание уже имеющихся клиентов. Среди задач, наиболее часто реализуемых с помощью ИИ, можно выделить управление поступающим контентом, осуществление простых процессов продаж, анализ информации о клиенте и формирование для него персональных предложений.

Инструменты искусственного интеллекта вместе с технологиями роботизированной автоматизации процессов (robotic process automation, RPA) захватывают сообщения и письма клиентов, идентифицируют их, распознают, извлекают необходимую и полезную информацию о клиенте и его запросе, верифицируют полученные данные с теми, которые уже есть у компании, и затем передают на обработку и принятие решения по конкретному вопросу.

Искусственный интеллект активно используется в простых процессах продаж на таких торговых площадках, как eBay и Amazon, в социальной сети Facebook и мессенджере WeChat. Yamato Transport, одна из крупнейших курьерских компаний Японии, использует чат-бот для планирования доставки и ответа на вопросы о том, где находятся посылки [24]. Domino's Pizza использует чат-бот, который принимает заказы на онлайн-доставку.

Для реализации поведенческого таргетинга в режиме реального времени на основе анализа транзакций клиентов для них формируются персональные предложения, опираясь на поведение покупателя и опыта продаж в компании. Например, подобные решения используются в Netflix, Amazon, Outbrain, Taboola [22].

Определение индивидуальной траектории взаимодействия с клиентом также может быть реализовано с помощью систем ИИ, в основе которых лежат технологии обработки естественного языка и машинного обучения. Например, компания Stitch Fix создала онлайн-магазин одежды, где покупателям предлагается определить их уникальный стиль (вместо выбора из предложенных шаблонных вариантов) и под него, с учетом индивидуальных особенностей человека, подобрать подходящие вещи [25].

Искусственный интеллект помогает усовершенствовать общение и повысить лояльность клиента, продать ему товар, за счет анализа его эмоционального состояния по голосу или тексту обращения. Подобные решения позволяют спрогнозировать поведение клиента, предугадать его желания и выстроить с ним взаимодействие наилучшим образом.

Роботизация и применение технологий искусственного интеллекта, как и многие другие инновации, позволяют удовлетворить клиентские потребности в качестве и скорости обработки их запросов [26–28]. Иногда даже удается превзойти ожидания покупателей и потребителей, сокращая их усилия на взаимодействие с компанией.

Исследование КПМГ [22], проведенное в 2020 году, показало, что за последнее время втрое увеличилось число потребителей, готовых использовать цифровые технологии (социальные сети, веб-чаты, мессенджеры) для взаимодействия с продавцами. По оценкам McKinsey [29], применение ИИ для более глубокой (по сравнению с традиционными решениями) аналитики данных позволяет компании увеличить ее ценность на 30–128%, а у розничных продавцов только в результате при-

менения технологий адаптации предложений под конкретного клиента продажи увеличиваются на 1–2%. Поэтому для укрепления своей конкурентоспособности, увеличения прибыли и улучшения клиентского опыта компании активно внедряют ИТ-решения, базирующиеся на методах машинного обучения.

Для непосредственного взаимодействия с клиентами компании начинают активно использовать искусственный интеллект. Исследование [21] показало, что применение скрытых чат-ботов (когда клиенты думают, что общаются с человеком) в четыре раза эффективнее неопытного продавца. Кроме того, подобные решения в некоторых ситуациях помогают превратить негативный опыт клиента в положительный, путем оперативного и прозрачного решения возникающих проблем.

2. Проблемы применения технологий искусственного интеллекта для взаимодействия с клиентами

Поведенческая экономика указывает на такую особенность человека, как формирование доверительного отношения к тем, кто нам нравится. Многие компании уже давно знают об этом и предпринимают соответствующие шаги, чтобы покупателю были симпатичны сотрудники, с которыми он общается. Однако пока остается нерешенным вопрос, как сделать так, чтобы покупателю было комфортно общаться с ИИ, который заменяет контактное лицо со стороны компании. По оценкам Cargemini, примерно две из пяти компаний, столкнувшихся с этическими проблемами использования ИИ, решили полностью отказаться от его применения [6].

Другими важными аспектами положительного клиентского опыта и повышения лояльности клиента являются сочувствие и персонализация, которые сложно реализуемы в случае применения технологий искусственного интеллекта. Теоретически, при проектировании алгоритма функционирования робота можно попытаться построить логику его взаимодействия с клиентом таким образом, чтобы он учитывал обстоятельства клиента и выказывал глубокое понимание его проблем, сомнений, опасений. Однако человек, общающийся, например, с ботом, скорее всего, не будет ощущать интерес к себе со стороны ИИ, не будет чувствовать себя ценным и уникальным, поскольку он будет понимать, что общается не с человеком, а с машиной. Клиен-

ты особенно не склонны взаимодействовать с роботом, если нужна субъективная оценка, помощь в подборе товара, соучастие и сочувствие [29–31].

Существует многочисленная категория клиентов, считающих, что компания, которая использует для взаимодействия с ними систему искусственного интеллекта, в некотором роде обманывает покупателя или потребителя услуг [24]. Это серьезно снижает доверие клиента к бренду. Так, после раскрытия информации о том, что в процессе продажи со стороны компании работает бот, а не человек, увеличивается частота прерывания текущего контакта, а количество покупок снижается почти на 80% [21].

Исследования, описанные в работе [28], показывают, что внедрение инновационных решений на базе ИИ у потенциальных клиентов и потребителей чаще вызывают отрицательную реакцию, чем положительную, особенно это сказывается на восприятии этической стороны репутации компании.

Клиентов сильно волнует конфиденциальность и безопасность взаимодействия с роботом [32]. При взаимодействии с живым человеком, в отличие от систем искусственного интеллекта, сохраняется ощущение, что разговор может остаться не зафиксированным, не вся информация, которая сообщается продавцу или менеджеру попадет в учетную систему и будет использована при дальнейших контактах или в каких-либо других целях. Кроме того, из-за непрозрачности алгоритмов функционирования ИИ, клиент чувствует себя неуверенно, передавая персональные данные о себе роботу.

В ходе применения искусственного интеллекта, например, при анализе клиентского опыта и взаимодействия с потребителем, возникает вопрос этического использования информации о клиенте. Принципы определения этичности поведения компании в данном случае также могут быть важным критерием прозрачности взаимодействия и честности организации. В желании предоставить клиентам персонализированное обслуживание многие фирмы злоупотребляют информацией о нем, что может негативным образом сказаться на потребительском опыте. Человеку может показаться, что продавец нарушает допустимые границы и лезет в его личную жизнь.

Обмен персональными данными между компаниями одной экосистемы в соответствии с законодательством Российской Федерации недопустим.

Однако передавать друг другу агрегированные, обезличенные данные в виде услуги возможно. Это позволяет использовать клиентский опыт участников экосистемы, не нарушая базовые этические нормы. Так, например, поступают компании «Мегафон» и «Mail.ru» [33]. Несмотря на то, что это не нарушает права клиентов, многие из них хотели бы знать о том, какие их персональные данные используются, кем и в каких целях.

Таким образом, решения на базе искусственного интеллекта воспринимаются многими людьми как инновационные, но с непонятным алгоритмом функционирования. Это влечет за собой ряд этических последствий, которые могут оказать негативное влияние и на клиентский опыт.

3. Этические принципы применения технологий искусственного интеллекта для взаимодействия с клиентами

Понятие «этичный» весьма многостороннее, оно может относиться как к процессу, так и к результату или ценности [18]. При рассмотрении этичности процесса речь будет идти о внутренних процедурах и действиях, которые реализует компания. Аспект этичности ценностей относится к набору параметров взаимодействия организации с покупателями и потребителями. В данном случае речь идет о прозрачности информационного взаимодействия, справедливости ценообразования, конфиденциальности персональных данных и т.п. Этичность результатов связана со свойствами выходных данных системы ИИ, отсутствием дискриминации, справедливостью и объективностью. Важно рассматривать этику ИИ со всех вышеперечисленных ракурсов, поскольку она затрагивает этические вопросы проектирования, разработки, внедрения и использования соответствующих технологий на практике.

Люди стали больше обращать внимание на то, насколько достойно ведут себя бренды в отношении своих этических и социальных обязательств. Многие исследователи [11, 13, 14] отмечают, что это способствует долгосрочному успеху компании в части сохранения лояльности клиентов.

Опасения, страх, непонимание механизмов функционирования систем, в которые встроены технологии ИИ, могут существенно снизить потенциальные положительные эффекты от их применения. Некоторые компании, активно тестирующие различные технологии ИИ (например, Walmart Inc.), обеспокоены отношением клиентов к робо-

там, с которыми они сталкиваются во время покупки [34]. Это в основном связано с этическим аспектом. В частности, имеют место следующие риски:

- ♦ возникновение у клиента ощущения, что за него принимают решение, то есть снижается возможность самореализации и обесцениваются способности человека к исследованию и выбору того продукта, который удовлетворит его потребности;
- ♦ непрозрачность зон ответственности за решения или выводы, принятые искусственным интеллектом. Например, если оператор колл-центра что-то советует звонящему покупателю, то ответственность за рекомендации лежит на конкретном сотруднике, человеке. Если же рекомендации дает робот, то кто будет нести за них ответственность?
- ♦ контроль за действиями искусственного интеллекта. Если система самообучаемая, то может возникать ситуация, когда выводы из анализируемой информации, поведение и предпринимаемые ИИ решения могут оказаться непрогнозируемыми и непредсказуемыми не только для пользователя, но и для разработчика.

Ведущие организации, работающие в области формирования этических норм и правил применения ИИ, сформулировали большое число разных подходов и инструментов для соблюдения вышеперечисленных принципов [1, 2, 4, 5]. Практически все они придерживаются схожих взглядов и совсем не противоречат друг другу, а только дополняют коллег.

В «белой книге» по этике ИИ Европейского института науки, СМИ и демократии (Atomium-European Institute for Science, Media and Democracy) приводятся пять принципов этики искусственного интеллекта: 1) содействие благополучию людей; 2) непричинение вреда (конфиденциальность, безопасность и «внимание к возможностям»); 3) автономия (право людей самостоятельно принимать решения); 4) справедливость (уважение интересов всех сторон, на которые могут повлиять действия системы с ИИ, отсутствие дискриминации, возможность устранения ошибок); 5) объяснимость (прозрачность логики искусственного интеллекта, подотчетность) [3]. Эти принципы представляют собой квинтэссенцию тех, которые изложены в кодексах, регламентах и прочих рекомендательных и нормативных документах, изданных экспертными и регулируемыми органами стран Европейского Союза.

Японское общество искусственного интеллекта (Japanese Society for Artificial Intelligence, JSAI) выделяет следующие этические принципы, кото-

рым должны следовать разработчики систем искусственного интеллекта [35]: 1) соблюдение прав человека и уважение культурного разнообразия; 2) соблюдение законов и правил, а также непричинение вреда другим; 3) уважение частной жизни; 4) справедливость; 5) безопасность; 6) добросовестность; 7) подотчетность и социальная ответственность; 8) саморазвитие и содействие пониманию ИИ обществом. Важно также отметить, что в отличие от европейских принципов ИИ, здесь особое внимание уделяется развитию ИИ таким образом, чтобы он соблюдал вышеперечисленные принципы в ходе своего функционирования.

Корпорация Google также сформулировала семь принципов искусственного интеллекта, которыми компания руководствуется при создании и использовании подобных технологий. К ним относятся [36]: 1) ИИ должен быть социально-полезным; 2) необходимо стремиться избегать несправедливого воздействия на людей; 3) применение передовой практики обеспечения безопасности; 4) ответственность за действия ИИ перед людьми; 5) обеспечение гарантий конфиденциальности, надлежащей прозрачности и контроля за использованием данных; 6) поддержание стандартов высокого мастерства; 7) ограничение использования потенциально вредоносных и оскорбительных программных продуктов. Особенностью данного свода является указание важности квалификации людей, создающих и

управляющих системами с ИИ.

В российском кодексе этики использования данных, который разработан по инициативе Ассоциации больших данных и Института развития интернета, в качестве основных принципов применения ИИ указано, что необходимо «базироваться на основополагающих принципах защиты прав и свобод человека, не допускать дискриминацию и причинение вреда», а также соблюдать российское и международное законодательство в области информационной безопасности и защиты данных от неправомерного использования [37].

Научно-исследовательский институт Carpgemini также сформулировал основные характеристики этичного ИИ. К ним относятся [6]: 1) этичность действий от проектирования до применения; 2) прозрачность; 3) объяснимость функционирования ИИ; 4) интерпретируемость результатов; 5) справедливость, отсутствие предвзятости; 6) возможность аудита.

В *таблице 1* приведены результаты сопоставления и сравнения сводов принципов, которые рекомендуются научными и экспертными сообществами [3, 6, 35], применяются ведущей ИТ-компанией Google [36], а также те, которые сформулированы в российском кодексе этики использования данных [37].

Таблица 1.

Сравнительная характеристика разных сводов принципов этики искусственного интеллекта

Принципы	Источник свода принципов				
	JSAI [35]	Atomium – EISMD [3]	Carpgemini [6]	Google [36]	АБД*) [37]
Равенство и справедливость	+	+	+	+	+
Польза, непричинение вреда	+	+		+	+
Уважение культурного разнообразия и плюрализма	+				+
Недискриминация и отсутствие стигматизации	+	+	+	+	+
Индивидуальная ответственность, подотчетность	+	+	+	+	
Автономия и согласие		+			
Конфиденциальность, уважение частной жизни	+	+		+	+
Защита данных и контроль их использования	+	+		+	+
Содействие пониманию ИИ обществом	+				
Добросовестность	+				
Социальная ответственность	+			+	
Высокое мастерство, саморазвитие	+			+	

*) Ассоциация больших данных и Институт развития интернета

Как отмечалось ранее, главная задача компании в этой области — это создать доверительные отношения с клиентом. На уровне компании важно соблюдать этические принципы, связанные с различными аспектами доверия. IBM среди них выделяет справедливость, надежность, прозрачность, подотчетность, объяснимость и выравнивание значений (соответствие правилам, бизнес-процессам, нормам, законам, этике и морали) [38].

Анализ рекомендаций по формированию этических принципов позволяет сделать вывод, что в качестве ключевых принципов этики искусственного интеллекта, которым необходимо соответствовать коммерческим компаниям, можно выделить две категории:

- ◆ категория «доверие»: справедливость, надежность, прозрачность, подотчетность, объяснимость;
- ◆ категория «соблюдение законодательства»: защита данных, контроль использования данных, обеспечение конфиденциальности.

Существенную роль в улучшении клиентского опыта также играют принятые бизнесом организационные меры по обеспечению соблюдения данных принципов. Примерами являются развитие совместного отраслевого регулирования и саморегулирования в области ИИ, разработка и утверждение моральных ценностей для компании, а также определение степени прозрачности функционирования искусственного интеллекта в системах управления рекламой, маркетингом, продажами и послепродажным обслуживанием.

4. Методы снижения негативного клиентского опыта взаимодействия с искусственным интеллектом

Этичное использование ИИ повышает лояльность клиента к компании, повышает доверие к ней и способствует росту объема продаж. Неэтичное поведение может привести к проявлению репутационных рисков, судебным разбирательствам и потере до 30% клиентов [6].

Для улучшения клиентского опыта взаимодействия с технологиями ИИ важно обеспечить соблюдение принципов вышеперечисленных категорий, а также дать возможность людям помогать компании совершенствовать алгоритмы и методы применения искусственного интеллекта. Таким образом компания будет демонстрировать приверженность базовым принципам этики ИИ.

Одним из главных способов повышения степени доверия при взаимодействии с искусственным интеллектом является информирование клиента о факте, что он общается с роботом. Важно отметить, что это информирование целесообразно осуществлять в конце процесса общения, когда человек сделал заказ или передал заявку / жалобу, которая поступила в обработку или даже решение по которой уже принято. Таким образом компания соблюдает этические нормы по обеспечению прозрачности применения ИИ в общении, а также позволяет клиенту получить положительный опыт общения с машиной. Если информировать об этом клиента в начале коммуникации, то возникает высокая вероятность прерывания данного контакта по инициативе клиента [21]. Из-за этого компания теряет лояльность клиента, а он, в свою очередь, не сможет сформировать свое мнение (возможно, положительное) о функционировании робота.

Другим инструментом является информирование клиента об алгоритме формирования предложений для него. Как правило, это делается по запросу, а не навязывается регулярно каждому человеку. Например, Facebook с помощью специальной опции «Почему я это вижу?» показывает критерии, которые сработали при подборе персонифицированной рекламы [22]. Также по запросу пользователя может создаваться информационный бюллетень, в котором представлены сведения о цифровой политике, используемых данных, цели сбора и обработки и данных, рисках, результатах проверки на соответствие этическим принципам компании. Так соблюдаются принципы объяснимости и подотчетности, сохраняется доверие пользователя к Facebook.

Если компания не раскрывает алгоритм формирования индивидуальной цены продукта или услуги, персональной скидки и рекламных предложений, то возникает ощущение обмана или дискриминации. На сегодняшний день уже есть много примеров подобного неэтичного применения ИИ. В качестве примера можно упомянуть страховую компанию Allstate в Мэриленде, формировавшую стоимость страховки в зависимости от платежеспособности клиента, которую она вычисляла по внешним источникам информации [18]. Такое скрытное поведение сегодня ужестораживает покупателя. Поэтому, аналогично раскрытию финансовой отчетности, публичные компании должны выкладывать в свободный доступ информацию о собираемых данных и политике их использования системами ИИ. Так будет реализовано соблюдение принципов прозрач-

ности и подотчетности, и в то же время обеспечена защита коммерческой тайны, поскольку детали не разглашаются, а вместо этого предоставляется информация общего характера.

Также целесообразным видится применение инструментов с открытым кодом, которые выявляют предвзятость и дискриминацию в алгоритмах искусственного интеллекта. В качестве примера можно привести AI Fairness 360 (AIF360), который представляет собой набор инструментов, позволяющий выявить и устранить предвзятость и смещение в моделях машинного обучения [39]. Подобные механизмы позволяют компании следовать этическим принципам, тактическим и стратегическим целям. Примером является практика выявления и снижения возрастной дискриминации при выдаче кредита в German Credit с помощью решения от IBM [40].

В регламенте Европейского Союза по защите данных (General Data Protection Regulation, GDPR) предусмотрена отчетность о применяемых алгоритмах ИИ для обеспечения соблюдения индивидуальных прав и системного совместного управления [41]. Это подразумевает проведение оценки воздействия на защиту данных (Data Protection Impact Assessment, DPIA). При ее выполнении в соответствии с концепцией «многоуровневого объяснения» [42] алгоритмических систем можно обеспечить реализацию клиентских прав на объяснение принципа функционирования ИИ, с которым они взаимодействуют.

Кроме того, важно доносить информацию об ИИ до клиентов таким образом, чтобы люди понимали ее, могли оценить важность и безопасность сбора сведений о них. Также рекомендуется сообщать клиентам и о рисках, которые сопутствуют функционированию систем на основе искусственного интеллекта, о возможном неправильном использовании или некорректной трактовке получаемых данных.

Отсюда следует и другой метод повышения доверия клиента — передача ему части функций по контролю и управлению областью функционирования системы искусственного интеллекта. Участник взаимодействия должен иметь возможность обратиться за помощью к реальному сотруднику компании или сообщить о некорректных действиях робота. При этом данная опция должна быть ему известна и легкодоступна. Клиенту можно делегировать полномочия по определению категорий персональных данных, доступных ИИ для обработки, а также перечня онлайн-сервисов, которые он хотел бы получать в ходе контакта с компанией. Например, в компании

«Билайн» [43] в качестве критерия этического поведения выделяется возможность клиента обозначить уровень персонализации взаимодействия. Это позволяет человеку ощущать себя управляющим условиями взаимодействия и чувствовать себя информационно защищенным.

Человек станет больше доверять искусственному интеллекту, если соответствующая система будет предоставлять актуальную статистику успешно выполненных работ, а также демонстрировать связь своих рекомендаций с целями человека, с которым она взаимодействует [44]. Здесь могут быть определены разные показатели производительности и результативности, например, количество клиентов, довольных обслуживанием конкретным роботом, доля успешно решенных проблем или средняя скорость обработки заявки.

Вышеописанные методы направлены на формирование доверительных, почти партнерских отношений между компанией и ее клиентами. Их реализация может не достигнуть желаемых эффектов, если эти действия будут лишь формальными, а не станут элементом корпоративной культуры фирмы. В обеспечении гибкости и открытости при выстраивании взаимоотношений с клиентами, особенно когда речь идет о применении искусственного интеллекта, важную роль играет цифровая культура, которая должна формироваться и развиваться в компании [45]. В организации должны быть созданы условия для развития этического мышления у ее сотрудников, чтобы они хотели поступать правильно и корректно на всех этапах создания и управления системой с искусственным интеллектом. Подобные мероприятия направлены главным образом на создание атмосферы эффективного и этического применения передовых информационных технологий. Но именно для укрепления доверия клиента нужно убедительно демонстрировать их во внешней среде. Целесообразным видится информирование общественности о своих принципах и методах развития корпоративной культуры, о целях, которые компания в этом направлении желает достигнуть и какие шаги она для этого предпринимает. Также важно показывать клиенту, что это не только декларация о намерениях, но и реальные действия. Например, что компания активно использует бенчмаркинг и ведет исследования в данной области, обучает сотрудников и создает внутренние центры компетенций по этике, а также принимает участие в различных сообществах и ассоциациях, занимающихся вопросами этики ИИ.

Несмотря на то, что целью внедрения подобных технологий является сокращение затрат на персонал, именно для формирования доверительных отношений на первом этапе внедрения ИИ, в рамках процесса послепродажного обслуживания сотрудник компании должен связаться с клиентом и поинтересоваться его опытом. Одной из задач такого контакта будет демонстрация клиенту намерений обеспечить ему комфорт и безопасность взаимодействия с ИИ, учитывая его сомнения и пожелания.

Применение предлагаемого перечня методов и подходов к улучшению взаимодействия с ИИ позволяет практически любой организации соответствовать ключевым этическим принципам и тем самым сохранить доверие клиента к компании.

Заключение

При активном использовании технологий ИИ в организации и реализации процессов продвижения, продажи и послепродажного обслуживания возникает проблема обеспечения положительного клиентского опыта. Это связано с тем, что многие люди боятся и не доверяют роботам и искусственному интеллекту. Для того, чтобы снизить негативное отношение к подобным системам, важно донести до клиента этичность применения этих решений, обеспечить прозрачность и подотчетность их функционирования.

В данной статье сформулированы основные направления применения технологий искусственного интеллекта в процессах взаимодействия с клиентом. Выявлена необходимость определения и формулирования этических принципов разработки, внедрения и применения ИИ на практике. В результате анализа рекомендаций международных и локальных экспертных комиссий, научных исследований и ведущих ИТ-компаний сформулированы две категории базовых принципов этики искусственного интеллекта, которым рекомендуется следовать коммерческим организациям.

Опираясь на данные принципы, научно-практические эксперименты и опыт компаний, активно использующих технологии ИИ, описаны ключевые методы улучшения клиентского опыта взаимодействия с подобными ИТ-решениями.

Представленные в данной статье рекомендации могут быть полезны организациям, которые планируют или уже применяют системы, функционирующие на базе ИИ. Полученные результаты также могут служить отправной точкой для дальнейших исследований. Например, интересным видится выявление изменения значимости тех или иных этических принципов в зависимости от отраслевой принадлежности компании, а также изучение влияния каждого из описанных методов на изменение доверия к ИИ клиентов российских компаний. ■

Литература

1. The Montreal declaration for a responsible development of artificial intelligence // Universite de Montreal, 2017. [Электронный ресурс]: <https://www.montrealdeclaration-responsibleai.com/> (дата обращения 20.08.2020).
2. Tenets of the partnership on AI // Partnership on AI, San Francisco. [Электронный ресурс]: <https://www.partnershiponai.org/tenets/> (дата обращения 21.08.2020).
3. AI 4People's ethical framework for a good AI society: Opportunities, risks, principles, and recommendations / L. Floridi [et al.] / AI4People, 2019. [Электронный ресурс]: <https://www.eismd.eu/wp-content/uploads/2019/03/AI4People%E2%80%99s-Ethical-Framework-for-a-Good-AI-Society.pdf> (дата обращения 21.08.2020).
4. What are the OECD Principles on AI? / OECD, 2019. [Электронный ресурс]: <https://www.oecd.org/going-digital/ai/principles> (дата обращения 27.08.2020).
5. Ethics guidelines for trustworthy AI / European Commission, 2019. [Электронный ресурс]: <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai> (дата обращения 29.08.2020).
6. Why addressing ethical questions in AI will benefit organizations / Capgemini Research Institute, 2019. [Электронный ресурс]: https://www.capgemini.com/wp-content/uploads/2019/08/AI-in-Ethics_Web.pdf (дата обращения 11.11.2020).
7. Allega P. Hype cycle for enterprise architecture, 2020 // Gartner, 2020. [Электронный ресурс]: <https://www.gartner.com/en/documents/3989875> (дата обращения 25.10.2020).
8. Shankar V. How artificial intelligence (AI) is reshaping retailing // Journal of Retailing. 2018. Vol. 94. No 4. P. vi–xi. DOI: 10.1016/S0022-4359(18)30076-9.
9. Huang M.-H., Rust R.T. Artificial intelligence in service // Journal of Service Research. 2018. Vol. 21. No 2. P. 155–172. DOI: 10.1177/1094670517752459.
10. Syam N., Sharma A. Waiting for a sales renaissance in the fourth industrial revolution: Machine learning and artificial intelligence in sales research and practice // Industrial Marketing Management. 2018. Vol. 69. P. 135–146. DOI: 10.1016/j.indmarman.2017.12.019.
11. Davenport T., Guha A., Grewal D., Bressgott T. How artificial intelligence will change the future of marketing // Journal of the Academy of Marketing Science. 2020. No 48. P. 24–42. DOI: 10.1007/s11747-019-00696-0.

12. Kaplan A., Haenlein M. Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence // *Business Horizons*. 2019. Vol. 62. No 1. P. 15–25. DOI: 10.1016/j.bushor.2018.08.004.
13. Fukukawa K., Balmer J.M., Gray E.R. Mapping the interface between corporate identity, ethics and corporate social responsibility // *Journal of Business Ethics*. 2007. No 76. P. 1–5. DOI: 10.1007/s10551-006-9277-0.
14. Stanaland A.J.S., Lwin M.O., Murphy P.E. Consumer perceptions of the antecedents and consequences of corporate social responsibility // *Journal of Business Ethics*. 2011. No 102. P. 47–55. DOI: 10.1007/s10551-011-0904-z.
15. Gray K. AI can be a troublesome teammate // *Harvard Business Review*. 2017. [Электронный ресурс]: <https://hbr.org/2017/07/ai-can-be-a-troublesome-teammate> (дата обращения 12.10.2020).
16. Lee J.D., See K.A. Trust in automation: Designing for appropriate reliance // *The Journal of the Human Factors and Ergonomics Society*. 2004. Vol. 46. No 1. P. 50–80. DOI: 10.1518/hfes.46.1.50_30392.
17. Customer experience in the new reality. Global Customer Experience Excellence research 2020: The COVID-19 special edition / KPMG, 2020. [Электронный ресурс]: <https://assets.kpmg/content/dam/kpmg/xx/pdf/2020/07/customer-experience-in-the-new-reality.pdf> (дата обращения 15.10.2020).
18. The state of AI ethics report / MAIEI, 2020. [Электронный ресурс]: <https://montrealethics.ai/the-state-of-ai-ethics-report-june-2020/> (дата обращения 29.08.2020).
19. Winfield A.F.T., Jirotko M. Ethical governance is essential to building trust in robotics and artificial intelligence systems // *Philosophical Transactions of the Royal Society A. Mathematical, physical and engineering sciences*. 2018. Vol. 376. No 2133. Article ID: 20180085. DOI: 10.1098/rsta.2018.0085.
20. С потребителем «на ты». 100 брендов с лучшим клиентским сервисом. Исследование КПМГ в России // KPMG, 2019. [Электронный ресурс]: <https://drive.google.com/file/d/1kdLf8yMxPi5N6ddkjUJrgrycIpppuq8o/view> (дата обращения 15.10.2020).
21. Luo X., Tong S., Fang Z., Qu Z. Frontiers: Machines vs. humans: The impact of artificial intelligence chatbot disclosure on customer purchases // *Marketing Science*. 2019. Vol. 38. No 6. P. 937–947. DOI: 10.1287/mksc.2019.1192.
22. Consumer choice and autonomy in the age of artificial intelligence and big data / Q. Andre [et al.] // *Customer Need and Solution*. 2018. No 5. P. 28–37. DOI: 10.1007/s40547-017-0085-8.
23. Service Encounter 2.0: An investigation into the roles of technology, employees and customers / B. Lariviere [et al.] // *Journal of Business Research*. 2017. No 79. P. 238–246. DOI: 10.1016/j.jbusres.2017.03.008.
24. Thompson C. May A.I. help you? Intelligent chatbots could automate away nearly all of our commercial interactions – for better or for worse // *New York Times*, 2018. [Электронный ресурс]: <https://www.nytimes.com/interactive/2018/11/14/magazine/tech-design-ai-chatbot.html> (дата обращения 28.09.2020).
25. Уилсон Дж., Доэрти П., Шукла П. Смелые одежды: как продавать вещи с помощью людей и искусственного интеллекта // *Harvard Business Review Россия*, 2017. [Электронный ресурс]: <https://hbr-russia.ru/innovatsii/tekhnologii/p18595/#ixzz4WWRr3huw> (дата обращения 04.11.2020).
26. Service research priorities in a rapidly changing context / A.L. Ostrom [et al.] // *Journal of Service Research*. 2015. Vol. 18. No 2. P. 127–159. DOI: 10.1177/1094670515576315.
27. Kim J., Kim K.H., Garrett T.C., Jung H. The contributions of firm innovativeness to customer value in purchasing behavior // *Journal of Product Innovation Management*. 2015. Vol. 32. No 2. P. 201–213. DOI: 10.1111/jpim.12173.
28. McLeay F., Osburg V.S., Yoganathan V., Patterson A. Replaced by a robot: Service implications in the age of the machine // *Journal of Service Research*. 2021. Vol. 24. No 1. P. 104–121. DOI: 10.1177/1094670520933354.
29. Notes from the AI frontier: Applications and value of deep learning / M. Chui [et al.] // McKinsey Global Institute discussion paper. 2018. [Электронный ресурс]: <https://www.mckinsey.com/featured-insights/artificial-intelligence/notes-from-the-ai-frontier-applications-and-value-of-deep-learning> (дата обращения 20.10.2020).
30. Froehlich A. Pros and cons of chatbots in the IT helpdesk // *InformationWeek*, 2018. [Электронный ресурс]: <https://www.informationweek.com/strategic-cio/it-strategy/pros-and-cons-of-chatbots-in-the-it-helpdesk/a/d-id/1332942> (дата обращения 20.10.2020).
31. Kestenbaum R. Conversational commerce is where online shopping was 15 years ago – Can it also become ubiquitous? // *Forbes*, 2018. [Электронный ресурс]: <https://www.forbes.com/sites/richardkestenbaum/2018/06/27/shopping-by-voice-is-small-now-but-it-has-huge-potential/?sh=3a41ebe337ac> (дата обращения: 27.09.2020).
32. Brave new world: Service robots in the frontline / J. Wirtz [et al.] // *Journal of Service Management*. 2018. Vol. 29. No 5. P. 907–931. DOI: 10.1108/JOSM-04-2018-0119.
33. Соболев А. Выручка от цифровых сервисов у «МегаФона» растет двузначными темпами // KPMG, 2020. [Электронный ресурс]: <https://mustread.kpmg.ru/interviews/vyruchka-ot-tsifrovyykh-servisov-u-megafona-rastet-dvuznachnymi-tempami/> (дата обращения: 28.10.2020).
34. Nassauer S. Walmart scraps plan to have robots scan shelves // *The Wall Street Journal*, 2020. [Электронный ресурс]: <https://www.wsj.com/articles/walmart-shelves-plan-to-have-robots-scan-shelves-11604345341> (дата обращения 04.11.2020).
35. The Japanese Society for Artificial Intelligence Ethical Guidelines / JSAI, 2017. [Электронный ресурс] <http://ai-elsi.org/wp-content/uploads/2017/05/JSAI-Ethical-Guidelines-1.pdf> (дата обращения 14.05.2021).
36. Pichai S. AI at Google: our principles // Google, 2018. [Электронный ресурс]: <https://blog.google/topics/ai/ai-principles/> (дата обращения 05.11.2020).
37. Кодекс этики использования данных / Ассоциация больших данных. Институт развития интернета. [Электронный ресурс]: <https://ac.gov.ru/files/content/25949/kodeks-etiki-pdf.pdf> (дата обращения 08.11.2020).

38. Experiences with improving the transparency of AI models and services / M. Hind [et al.] // Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (CHI EA '20). Honolulu, USA, 25–30 April 2020. P. 1–8. DOI: 10.1145/3334480.3383051.
39. AI Fairness 360 / IBM Developer Staff, 2018. [Электронный ресурс]: <https://developer.ibm.com/technologies/artificial-intelligence/projects/ai-fairness-360/> (дата обращения 08.11.2020).
40. Detecting and mitigating age bias on credit decisions // Jupyter nbviewer. [Электронный ресурс]: https://nbviewer.jupyter.org/github/IBM/AIF360/blob/master/examples/tutorial_credit_scoring.ipynb (дата обращения 05.11.2020).
41. О GDPR на русском. Информация об Общем регламенте по защите данных. [Электронный ресурс]: <https://ogdpr.eu/> (дата обращения 08.11.2020).
42. Kaminski M.E., Malgieri G. Algorithmic impact assessments under the GDPR: producing multi-layered explanations // International Data Privacy Law. 2020. ipaa020. DOI: 10.1093/idpl/ipaa020.
43. Елаева М. Мы будем перестраивать все наши бизнес-процессы «от клиента» / КПМГ, 2020. [Электронный ресурс]: <https://mustread.kpmg.ru/interviews/my-budem-perestraivat-vse-nashi-biznes-protsessy-ot-klienta/> (дата обращения 01.10.2020).
44. Lara F., Deckers J. Artificial intelligence as a Socratic assistant for moral enhancement // Neuroethics. 2020. No 13. P. 275–287. DOI: 10.1007/s12152-019-09401-y.
45. Долганова О.И., Деева Е.А. Готовность компании к цифровым преобразованиям: проблемы и диагностика // Бизнес-информатика. 2019. Т. 13. № 2. С. 59–72. DOI: 10.17323/1998-0663.2019.2.59.72.

Об авторе

Долганова Ольга Игоревна

кандидат экономических наук;

доцент Департамента бизнес-информатики, Финансовый университет при Правительстве Российской Федерации, 105187, г. Москва, ул. Щербаковская, д. 38;

E-mail: oidolganova@fa.ru

ORCID: 0000-0001-6060-5421

Improving customer experience with artificial intelligence by adhering to ethical principles

Olga I. Dolganova

E-mail: oidolganova@fa.ru

Financial University under the Government of the Russian Federation
Address: 38, Scherbakovskaya Street, Moscow 105187, Russia

Abstract

The intensive development and application of artificial intelligence technologies in organizing interaction with clients is accompanied by such difficulties as: the client's unwillingness to communicate with the robot, distrust, fear, negative experience of the clients. Such problems can be solved by adhering to ethical principles of using artificial intelligence. In scientific and practical research on this topic, there are many general recommendations that are difficult to apply in practice, or, on the contrary, that describe the methods for solving a highly specialized technical or management problem. The purpose of this article is to determine the ethical principles and methods, the observance and implementation of which would increase confidence in artificial intelligence systems among client of a particular organization. As a result of the analysis and synthesis of the scientific and practical investigations, as well as the empirical experience of Russian and foreign companies, the main areas of application of artificial intelligence technologies affecting the customer experience were identified. The ethical principles recommended to be followed by business have been formulated and systematized. The main methods have been also identified to enable implementation of these principles in practice, and so to reduce the negative effects of customer interaction with artificial intelligence and increase their confidence in the company.

Key words: ethics; artificial intelligence (AI); ethical AI; customer experience; ethical principles; machine learning; trust; robot.

Citation: Dolganova O.I. (2021) Improving customer experience with artificial intelligence by adhering to ethical principles. *Business Informatics*, vol. 15, no 2, pp. 34–46. DOI: 10.17323/2587-814X.2021.2.34.46

References

1. Universite de Montreal (2017) *The Montreal declaration for a responsible development of artificial intelligence*. Available at: <https://www.montreal-declaration-responsibleai.com/> (accessed 20 August 2020).
2. Partnership on AI (2020) *Tenets of the partnership on AI*. Available at: <https://www.partnershiponai.org/tenets/> (accessed 21 August 2020).
3. Floridi L., Beltrametti M., Burri T., Chatila R., Chazerand P., Dignum V., Madelin R., Luetge C., Pagallo U., Rossi F., Schafer B., Valcke P., Vayena E. (2020) *AI 4People's ethical framework for a good AI society: Opportunities, risks, principles, and recommendations*. Available at: <https://www.eismd.eu/wp-content/uploads/2019/03/AI4People%E2%80%99s-Ethical-Framework-for-a-Good-AI-Society.pdf> (accessed 21 August 2020).
4. OECD (2019) *What are the OECD Principles on AI?* Available at: <https://www.oecd.org/going-digital/ai/principles> (accessed 27 August 2020).
5. European Commission (2019) *Ethics guidelines for trustworthy AI*. Available at: <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai> (accessed 29 August 2020).
6. Capgemini Research Institute (2019) *Why addressing ethical questions in AI will benefit organizations*. Available at: https://www.capgemini.com/wp-content/uploads/2019/08/AI-in-Ethics_Web.pdf (accessed 11 November 2020).
7. Allega P. (2020) Hype cycle for enterprise architecture, 2020. *Gartner*. Available at: <https://www.gartner.com/en/documents/3989875> (accessed 25 October 2020).
8. Shankar V. (2018) How artificial intelligence (AI) is reshaping retailing. *Journal of Retailing*, vol. 94, no 4, pp. vi–xi. DOI: 10.1016/S0022-4359(18)30076-9.
9. Huang M.-H., Rust R.T. (2018) Artificial intelligence in service. *Journal of Service Research*, vol. 21, no 2, pp. 155–172. DOI: 10.1177/1094670517752459.
10. Syam N., Sharma A. (2018) Waiting for a sales renaissance in the fourth industrial revolution: Machine learning and artificial intelligence in sales research and practice. *Industrial Marketing Management*, vol. 69, pp. 135–146. DOI: 10.1016/j.indmarman.2017.12.019.
11. Davenport T., Guha A., Grewal D., Bressgott T. (2020) How artificial intelligence will change the future of marketing. *Journal of the Academy of Marketing Science*, no 48, pp. 24–42. DOI: 10.1007/s11747-019-00696-0.
12. Kaplan A., Haenlein M. (2019) Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, vol. 62, no 1, pp. 15–25. DOI: 10.1016/j.bushor.2018.08.004.
13. Fukukawa K., Balmer J.M., Gray E.R. (2007) Mapping the interface between corporate identity, ethics and corporate social responsibility. *Journal of Business Ethics*, no 76, pp. 1–5. DOI: 10.1007/s10551-006-9277-0.
14. Stanaland A.J.S., Lwin M.O., Murphy P.E. (2011) Consumer perceptions of the antecedents and consequences of corporate social responsibility. *Journal of Business Ethics*, no 102, pp. 47–55. DOI: 10.1007/s10551-011-0904-z.
15. Gray K. (2017) AI can be a troublesome teammate. *Harvard Business Review*. Available at: <https://hbr.org/2017/07/ai-can-be-a-troublesome-teammate> (accessed 12 October 2020).
16. Lee J.D., See K.A. (2004) Trust in automation: Designing for appropriate reliance. *The Journal of the Human Factors and Ergonomics Society*, vol. 46, no 1, pp. 50–80. DOI: 10.1518/hfes.46.1.50_30392.
17. KPMG (2020) *Customer experience in the new reality. Global Customer Experience Excellence research 2020: The COVID-19 special edition*. Available at: <https://assets.kpmg/content/dam/kpmg/xx/pdf/2020/07/customer-experience-in-the-new-reality.pdf> (accessed 15 October 2020).
18. MAIEI (2020) *The state of AI ethics report*. Available at: <https://montrealaiethics.ai/the-state-of-ai-ethics-report-june-2020/> (accessed 29 August 2020).
19. Winfield A.F.T., Jiroka M. (2018) Ethical governance is essential to building trust in robotics and artificial intelligence systems. *Philosophical Transactions of the Royal Society A. Mathematical, physical and engineering sciences*, vol. 376, no 2133, article ID: 20180085. DOI: 10.1098/rsta.2018.0085.
20. KPMG (2019) *With the consumer "on thou". 100 brands with the best customer service. KPMG research in Russia*. Available at: <https://drive.google.com/file/d/1kdLf8yMxPi5N6ddkUJrgryclpppuq8o/view> (accessed 15 October 2020) (in Russian).
21. Luo X., Tong S., Fang Z., Qu Z. (2019) Frontiers: Machines vs. humans: The impact of artificial intelligence chatbot disclosure on customer purchases. *Marketing Science*, vol. 38, no 6, pp. 937–947. DOI: 10.1287/mksc.2019.1192.
22. Andre Q., Carmon Z., Wertenbroch K., Crum A., Frank D., Goldstein W., Huber J., van Boven L., Weber B., Yang H. (2018) Consumer choice and autonomy in the age of artificial intelligence and big data. *Customer Need and Solution*, no 5, pp. 28–37. DOI: 10.1007/s40547-017-0085-8.
23. Lariviere B., Bowen D., Andreassen T.W., Kunz W., Sirianni N.J., Voss C., Wunderlich N.V., De Keyser A. (2017) Service Encounter 2.0: An investigation into the roles of technology, employees and customers. *Journal of Business Research*, no 79, pp. 238–246. DOI: 10.1016/j.jbusres.2017.03.008.
24. Thompson C. (2018) May A.I. help you? Intelligent chatbots could automate away nearly all of our commercial interactions – for better or for worse. *New York Times*. Available at: <https://www.nytimes.com/interactive/2018/11/14/magazine/tech-design-ai-chatbot.html> (accessed 28 September 2020).

25. Wilson J., Daugherty P., Shukla P. (2017) Bold clothes: How to sell things with the help of people and artificial intelligence. *Harvard Business Review Russia*. Available at: <https://hbr-russia.ru/innovatsii/tehnologii/p18595/#ixzz4WWR3huw> (accessed 04 November 2020) (in Russian).
26. Ostrom A.L., Parasuraman A., Bowen D.E., Patrício L., Voss C.A. (2015) Service research priorities in a rapidly changing context. *Journal of Service Research*, vol. 18, no 2, pp. 127–159. DOI: 10.1177/1094670515576315.
27. Kim J., Kim K.H., Garrett T.C., Jung H. (2015) The contributions of firm innovativeness to customer value in purchasing behavior. *Journal of Product Innovation Management*, vol. 32, no 2, pp. 201–213. DOI: 10.1111/jpim.12173.
28. McLeay F., Osburg V.S., Yoganathan V., Patterson A. (2021) Replaced by a robot: Service implications in the age of the machine. *Journal of Service Research*, vol. 24, no 1, pp. 104–121. DOI: 10.1177/1094670520933354.
29. Chui M., Manyika J., Miremadi M., Henke N., Chung R., Nel P., Malhotra S. (2018) *Notes from the AI frontier: Applications and value of deep learning*. McKinsey Global Institute discussion paper. Available at: <https://www.mckinsey.com/featured-insights/artificial-intelligence/notes-from-the-ai-frontier-applications-and-value-of-deep-learning> (accessed 20 October 2020).
30. Froehlich A. (2018) Pros and cons of chatbots in the IT helpdesk. *InformationWeek*. Available at: <https://www.informationweek.com/strategic-cio/it-strategy/pros-and-cons-of-chatbots-in-the-it-helpdesk/a/d-id/1332942> (accessed 20 October 2020).
31. Kestenbaum R. (2018) Conversational commerce is where online shopping was 15 years ago – Can it also become ubiquitous? *Forbes*. Available at: <https://www.forbes.com/sites/richardkestenbaum/2018/06/27/shopping-by-voice-is-small-now-but-it-has-huge-potential/?sh=3a41ebe337ac> (дата обращения: 27.09.2020).
32. Wirtz J., Patterson P.G., Kunz W.H., Gruber T., Lu V.N., Paluch S., Martins A. (2018) Brave new world: Service robots in the frontline. *Journal of Service Management*, vol. 29, no 5, pp. 907–931. DOI: 10.1108/JOSM-04-2018-0119.
33. Sobolev A. (2020) Revenue from digital services at MegaFon is growing at a double-digit rate. *KPMG*. Available at: <https://mustread.kpmg.ru/interviews/vyruchka-ot-tsifrovyykh-servisov-u-megafona-rastet-dvuznachnymi-tempami/> (accessed 28 October 2020) (in Russian).
34. Nassauer S. (2020) Walmart scraps plan to have robots scan shelves. *The Wall Street Journal*. Available at: <https://www.wsj.com/articles/walmart-shelves-plan-to-have-robots-scan-shelves-11604345341> (accessed 04 November 2020).
35. JSAI (2017) *The Japanese Society for Artificial Intelligence Ethical Guidelines*. Available at: <http://ai-elsi.org/wp-content/uploads/2017/05/JSAI-Ethical-Guidelines-1.pdf> (accessed 14 May 2021).
36. Pichai S. (2018) AI at Google: our principles. *Google*. Available at: <https://blog.google/topics/ai/ai-principles/> (accessed 05 November 2020).
37. Big Data Association. Institute of internet development (2019) *Data usage code of ethics*. Available at: <https://ac.gov.ru/files/content/25949/kodeks-etiki-pdf.pdf> (accessed 08 November 2020) (in Russian).
38. Hind M., Houde S., Martino J., Mojsilovic A., Piorowski D., Richards J., Varshney K.R. (2020) Experiences with improving the transparency of AI models and services. Extended Abstracts of the *2020 CHI Conference on Human Factors in Computing Systems (CHI EA '20)*. Honolulu, USA, 25–30 April 2020. P. 1–8. DOI: 10.1145/3334480.3383051.
39. IBM Developer Staff (2018) *AI Fairness 360*. Available at: <https://developer.ibm.com/technologies/artificial-intelligence/projects/ai-fairness-360/> (accessed 08 November 2020).
40. Jupyter nbviewer (2020) *Detecting and mitigating age bias on credit decisions*. Available at: https://nbviewer.jupyter.org/github/IBM/AIF360/blob/master/examples/tutorial_credit_scoring.ipynb (accessed 05 November 2020).
41. *About GDPR in Russian. Information about the General Data Protection Regulation*. Available at: <https://ogdpr.eu/> (accessed 08 November 2020) (in Russian).
42. Kaminski M.E., Malgieri G. (2020) Algorithmic impact assessments under the GDPR: producing multi-layered explanations. *International Data Privacy Law*, ipaa020. DOI: 10.1093/idpl/ipaa020.
43. Elaeva M. (2020) We will rebuild all our business processes “from the client”. *KPMG*. Available at: <https://mustread.kpmg.ru/interviews/my-budem-perestraitvat-vse-nashi-biznes-protsessy-ot-klienta/> (accessed 01 October 2020) (in Russian).
44. Lara F., Deckers J. (2020) Artificial intelligence as a Socratic assistant for moral enhancement. *Neuroethics*, no 13, pp. 275–287. DOI: 10.1007/s12152-019-09401-y.
45. Dolganova O.I., Deeva E.A. (2019) Company readiness for digital transformations: problems and diagnosis. *Business Informatics*, vol. 13, no 2, pp. 59–72. DOI: 10.17323/1998-0663.2019.2.59.72.

About the author

Olga I. Dolganova

Cand. Sci. (Econ.);

Associate Professor, Department of Business Informatics, Financial University under the Government of the Russian Federation, 38, Scherbakovskaya Street, Moscow 105187, Russia;

E-mail: oidolganova@fa.ru

ORCID: 0000-0001-6060-5421

[DOI: 10.17323/2587-814X.2021.2.47.59](https://doi.org/10.17323/2587-814X.2021.2.47.59)

Системы интернета вещей в процессе мультидисциплинарной подготовки кадров для цифровой экономики и их проектирование

В.А. Комаров^{a,b} 

E-mail: VKomarov@iss-reshetnev.ru

А.В. Сарафанов^c 

E-mail: Sarafanov@i-teco.ru

^a АО «Информационные спутниковые системы» имени академика М.Ф. Решетнева»
Адрес: 662972, г. Железногорск, ул. Ленина, д. 52

^b Сибирский федеральный университет
Адрес: 660041, г. Красноярск, пр. Свободный, д. 79

^c ООО «Витте Консалтинг» (ГК «Ай-Теко»)
Адрес: 117036, Москва, ул. Кедрова, д. 15

Аннотация

В условиях цифровизации наиболее наукоемких отраслей экономики большое значение имеет развитие отраслевой системы подготовки кадров в области электронного приборостроения. К ключевым направлениям ее развития с применением информационно-коммуникационных технологий относятся разработка и совершенствование технологического базиса для подготовки и переподготовки кадров по инженерным образовательным программам. Одним из элементов такого базиса является сервис многопользовательского удаленного доступа к высокотехнологичному экспериментальному оборудованию на основе систем интернета вещей (internet of things, IoT), получивший название «лаборатория как сервис» (laboratory as a service). В рамках данного сервиса актуальной проблемой является повышение функциональной насыщенности автоматизированных макетов, стендов и установок. Данная проблема в настоящее время характеризуется малочисленностью научных исследований. Поэтому цель настоящей работы состоит в расширении областей экспериментальных исследований, выполняемых в режиме многопользовательского удаленного доступа на основе специализированных IoT-систем. В результате разработан метод мультидисциплинарного применения специализированных IoT-систем, заключающийся в технической реализации возможности дополнительных исследований, включая исследования технологий, лежащих в основе как многопользовательских распределенных измерительно-управляющих систем, так и IoT-систем в целом, исследования технологий, используемых при их сквозном автоматизированном проектировании, а также исследования совместного взаимодействия нескольких территориально распределенных автоматизированных макетов, стендов и установок, реализованных на основе четырехуровневой архитектуры IoT. Также разработана методология проектирования многопользовательских распределенных измерительно-управляющих систем как специализированных IoT-систем, ориентированная на решение мультидисциплинарных исследовательских задач в интерактивном диалоговом режиме на основе единичных экземпляров экспериментального оборудования. Данная методология обеспечивает организационное, техническое и методическое сопровождение процесса создания таких систем с заданными целевыми характеристиками. В целом разработанные метод и методология открывают возможности реализации на системной основе базовых принципов концепции «Образование 4.0» при подготовке и переподготовке инженерных кадров в области электронного приборостроения.

Ключевые слова: интернет вещей; IoT-система; лаборатория как сервис; технические средства обучения; многопользовательские распределенные измерительно-управляющие системы; методология проектирования.

Цитирование: Комаров В.А., Сарафанов А.В. Системы интернета вещей в процессе мультидисциплинарной подготовки кадров для цифровой экономики и их проектирование // Бизнес-информатика. 2021. Т. 15. № 2.

С. 47–59. DOI: 10.17323/2587-814X.2021.2.47.59

Введение

Ускоренные темпы цифровизации различных сфер человеческой деятельности отражаются, в первую очередь, на содержательной, методической и технологической составляющих отраслевой системы подготовки кадров (ОСПК). При этом особенностью текущей ситуации в мире, накладывающей ряд ограничений на традиционные подходы к образованию, является пандемия COVID-19 [1]. В этих условиях повышается актуальность следующих современных тенденций в ОСПК:

- ♦ применение на системной основе в различных видах учебных занятий информационно-коммуникационных технологий (ИКТ);
- ♦ усиление акцента на персонализацию обучения, внедрение гибкого графика обучения и отказ от жесткой привязки обучаемых к конкретным локациям;
- ♦ модернизация традиционной инфраструктуры образовательных учреждений – внедрение цифровых библиотечных систем, широкое использование ИКТ-инструментов и специализированных систем для одновременной организации лекционных занятий во множестве различных локаций для онлайн и оффлайн-режимов, применение кибер-физических систем, а также систем интернета вещей (IoT-систем), в частности, в качестве средств для проведения экспериментов в специализированных распределенных лабораториях;
- ♦ организация обучения с применением ИКТ как непрерывного процесса, сопровождающего человека на протяжении всей его жизни [2–10].

В рамках стратегии «Образование 4.0» выделяют следующие ключевые аспекты [11]:

- ♦ образование по требованию (education on demand);
- ♦ персонализация образования (personalization of education);
- ♦ обучение в удобное время в любом месте (learning at a convenient time anywhere);
- ♦ обучение на протяжении всей жизни (life-long learning).

Практическая реализация перечисленных аспектов опирается на разработку и совершенствование технологического базиса с применением ИКТ для подготовки и переподготовки кадров по инженерным образовательным программам в области электронного приборостроения. Так, при подготовке и переподготовке инженерных кадров для цифровой экономики реализация таких аспектов как «обучение в удобное время в любом месте» и «персонализация образования» в рамках цифровых образовательных сред базируются на усовершенствовании сервиса доступа к высокотехнологичному экспериментальному оборудованию и современным специализированным программным средствам – «лаборатория как сервис» (laboratory as a service) [12–14]. Именно такой сервис позволяет обучающимся выполнять натурные и вычислительные экспериментальные исследования в режиме удаленного доступа по сети интернет из любой локации.

Актуальность развития данного сервиса связана с повышением его эффективности. К основным направлениям повышения эффективности сервиса относятся повышение пропускной способности лабораторного оборудования, входящего в состав удаленных лабораторий, а также его функциональной насыщенности в части экспериментальных исследований как в разрезе мультидисциплинарности, так и приобретаемых в процессе обучения практик. Пути решения проблемы в части повышения пропускной способности высокотехнологичного экспериментального оборудования авторами данной статьи уже изложены в ряде работ [14–15], а подход к решению проблемы повышения функциональной насыщенности экспериментального оборудования является предметом данной статьи.

1. Метод мультидисциплинарного применения экспериментального оборудования с удаленным доступом

Основой технического обеспечения удаленных лабораторий являются технические средства обучения (ТСО) [16], разработанные с применением сквозной цифровой технологии, в том числе технологии ин-

тернета вещей (internet of things, IoT) [2–4, 17–19]. К таким ТСО в числе других относятся многопользовательские распределенные измерительно-управляющие системы (МРИУС), реализующие технологию эксплуатации высокотехнологичного оборудования (измерительных средств, средств вычислительной техники, автоматизированных лабораторных макетов/стендов/установок) на основе концепции мультиарендности [14]. Схема интеграции МРИУС в опорную четырехуровневую архитектуру IoT приведена на *рисунке 1* [20, 21].

Процесс создания образцов МРИУС как соответствующих IoT-систем базируется на интеграции целого пула цифровых технологий, которые, в свою очередь, могут стать предметом отдельных экспериментальных исследований в дополнение к целевым исследованиям в предметной области, определяемой соответствующими автоматизированными макетами/стендами/установками (*рисунок 1*). Кроме того, передовые технологии проектирования таких IoT-систем также могут служить отдельным предметом лабораторных исследований в разрезе сквозных циклов автоматизированного проектирования как отдельных элементов, так и данных систем в целом. Таким образом, рассматриваемые современные ТСО в виде МРИУС (как соответствующих IoT-систем) характеризуются следующими направлениями лабораторных исследований:

- ◆ проведение натурных экспериментальных исследований в целевой предметной области, определяемой конкретными АЛМ или их совокупностью (целевое назначение МРИУС);
- ◆ исследование технологий, лежащих в основе функционирования МРИУС (дополнение к целевому назначению);
- ◆ исследование технологий, используемых при проектировании МРИУС (дополнение к целевому назначению).

Каждое из данных направлений характеризуется определенной областью исследований и своим набором развиваемых в процессе обучения практик, приведенных в *таблице 1*. Для более подробного анализа составляющих эксперимента в *таблице 1* вводятся следующие уточнения:

- ◆ «учебным экспериментом» называются натурные экспериментальные исследования, предусмотренные соответствующими учебными курсами и/или программами дисциплин;
- ◆ к «научному эксперименту» отнесены натурные экспериментальные исследования, направленные

на получение информации о новых, ранее не исследованных процессах и/или явлениях (в рамках учебного процесса).

В своей совокупности перечисленные направления позволяют применять соответствующие МРИУС как ТСО для разработки учебных курсов, носящих мультидисциплинарный характер. Это, в свою очередь, позволяет развивать у обучающихся в рамках профессиональных компетенций такие навыки, как «системные способности», «когнитивные способности» и «решение сложных проблем», которые являются достаточно востребованными в основных секторах цифровой экономики. Кроме того, ряд направлений позволяет реализовывать контекстное обучение, например, в такой области исследования, как технология сквозного автоматизированного проектирования.

Как известно, возрастающая функциональная сложность и стоимость устройств, комплексов и систем затрудняют надлежащее оснащение учебно-исследовательских лабораторий их современными образцами, как объектами соответствующих экспериментальных исследований [22–24]. При этом использование свойства МРИУС, как IoT-системы (*рисунок 1б*), а именно совместного взаимодействия нескольких МРИУС, обеспечивает возможность исследования на их основе нескольких вариантов реализации устройств/комплексов/систем в целом. Организация таких исследований заключается в следующем. Соответствующие АЛМ, реализующие вариативность исполнения функциональных узлов/устройств/комплексов, могут размещаться как в одной локации и предусматривать возможность их соединения на физическом уровне для организации их взаимодействия, так и могут быть распределены территориально по нескольким локациям. В данном случае представляется возможным их «собрать» (эмулировать процесс их совместного функционирования) в составе общего устройства/комплекса/системы на **сервисном уровне** (*рисунок 1*) посредством организации взаимодействия М2М на **коммуникационном уровне**. Такой вид взаимодействия реализуется посредством создания цифровых «слепков» соответствующих сигналов (оцифровки их фрагментов) с последующим обменом ими между МРИУС различных локаций. В целях реализации такого взаимодействия в определенных образцах МРИУС предусматриваются технические устройства (средства) для записи цифровых отсчетов соответствующих сигналов и/или формирования сигналов по их оцифрованным фрагментам.

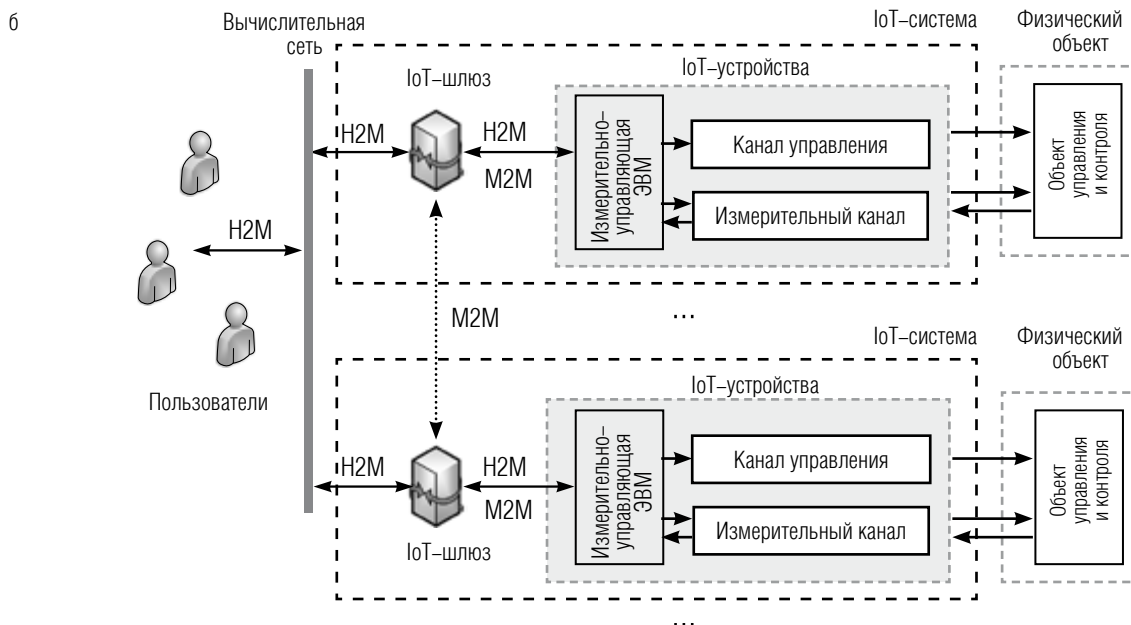
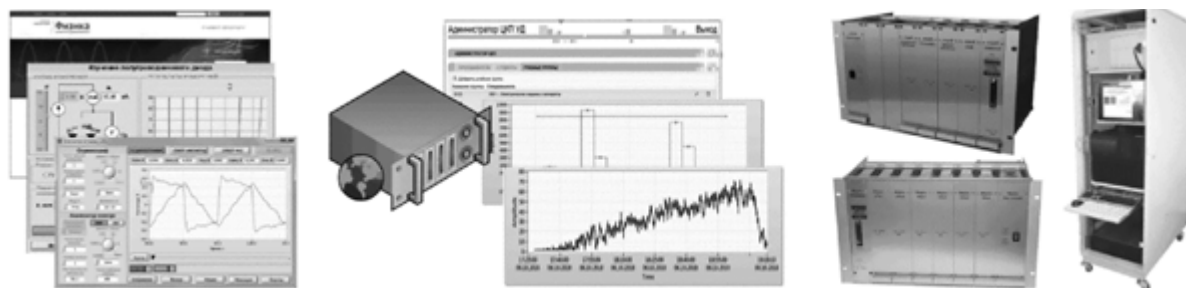
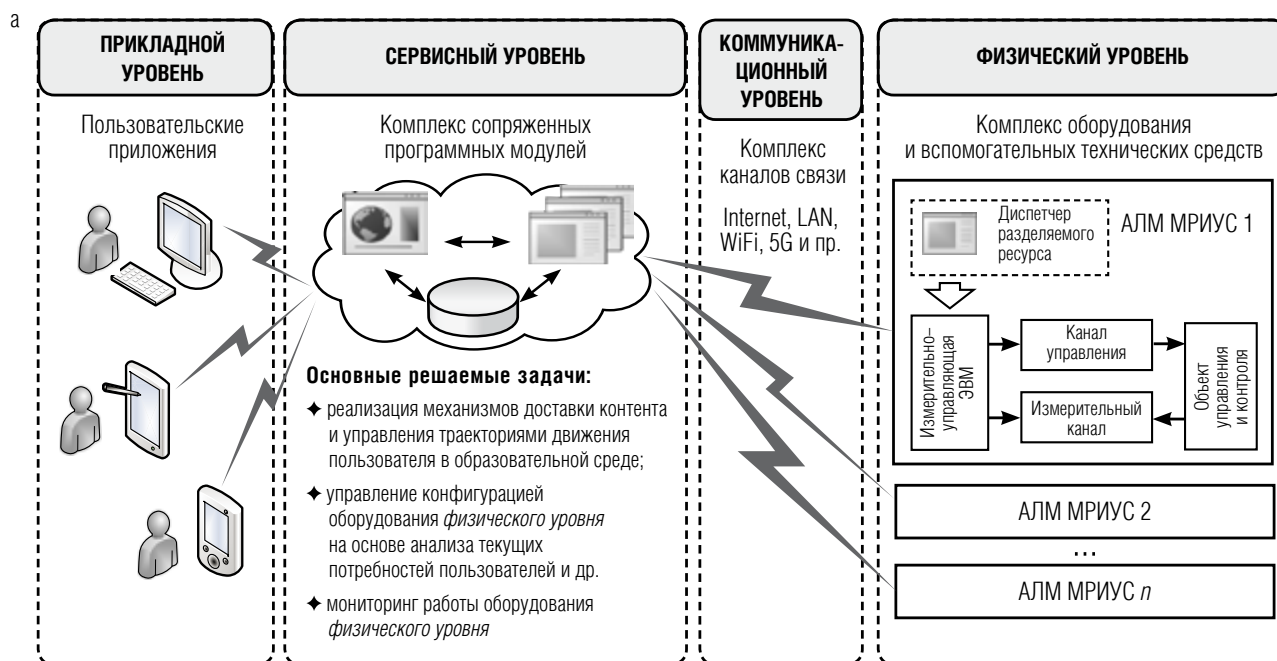


Рис. 1. МРИУС как IoT-система с четырехуровневой архитектурой:

а – уровни IoT; б – МРИУС как IoT-система;

АЛМ – автоматизированный лабораторный макет;

H2M, M2M – принципы организации взаимодействия (Human-to-Machine, Machine-to-Machine)

Таблица 1.

Направления мультидисциплинарных исследований на основе МРИУС

№ п/п	Направление исследований	Область исследований	Виды экспериментальных исследований	Развиваемые практики в области разработки/исследований
1.	Предметная область, определяемая АТМ, а также свойствами МРИУС как IoT-систем в целом	1.1 Физические процессы или явления в отдельных функциональных узлах или устройствах РЭС, а также устройствах комплексов и системах РЭС. 1.2 Методы организации внешнего (сетевого) взаимодействия между элементами IoT-систем по принципу M2M	Учебные / научные	1. Применение современного ИКТ-инструментария для анализа процессов, протекающих в РЭС. 2. Применение принципов построения IoT-систем как одного из перспективных направлений в приборостроении. 3. Применение МРИУС для автоматизации учебного и научного эксперимента, как сегмента современной ЦОС*). 4. Применение методов электронной педагогики, повышающих качество преподавания технических дисциплин*). 5. Управление индивидуальными образовательными траекториями в ЦОС*).
2.	Технологии, лежащие в основе функционирования МРИУС	2.1 Средства дистанционного управления объектами и технологическими процессами по вычислительным сетям	Учебные	Применение технологий передачи данных в IoT-системах
		2.2 Технологии и протоколы передачи данных по цифровым каналам	Учебные / научные	
		2.3 Статистические закономерности передачи трафика по цифровым каналам	Научные	
		2.4 Процессы функционирования и диспетчеризации запросов в МРИУС как в интерактивных диалоговых системах (системах массового обслуживания, СМО)	Учебные / научные	
		2.5 Поведение пользователей в эргатических (человеко-машинных) системах	Научные	
3.	Технологии проектирования МРИУС	2.6 Технологии функционального тестирования графического интерфейса пользователя (ГИП)	Учебные / научные	Применение методов функционального тестирования ГИП для человеко-машинных систем (систем управления объектами, виртуальных измерительных систем и др.)
		2.7 Методы обработки результатов экспериментальных исследований	Учебные	Применение специализированных программных средств для цифровой, статистической и прочей обработки результатов экспериментальных исследований
		2.8 Компьютерные измерительные технологии	Учебные	Применение компьютерных измерительных технологий при разработке IoT-систем
		3.1 Технологии сквозного автоматизированного проектирования / «Цифровой двойник»	Учебные	3.1 Применение методов математического моделирования в процессе проектирования IoT-систем и кибер-физических систем в целом.
				3.2 Применение методов идентификации и верификации математических моделей в процессе исследований электрических характеристик РЭС.
				3.3 Применение методов разработки современной электронной конструкторской документации.
				3.4 Изучение особенностей методического обеспечения для сквозного автоматизированного проектирования РЭС на базе САЕ/CAD/CAM/PLM-систем (конвейерное, циклическое и итерационное решение проектных задач на базе специализированных инженерных методик)
		3.2 Технологии и средства разработки прикладного ПО (frontend / backend) специализированных IoT-систем	Учебные	Применение технологий и инструментов разработки прикладного ПО для автоматизированных систем (систем управления объектами, виртуальных измерительных систем и др.)

*) Развиваемые практики в процессе переподготовки педагогических кадров

В качестве примера такой системы, реализующей данный принцип взаимодействия, на *рисунке 2* приведен пример структурной схемы IoT-суперсистемы для обеспечения процесса подготовки и переподготовки специалистов для авиакосмической отрасли. В данной системе на основе комплексирования взаимодействий H2M и M2M реализуется макет системы спутниковой связи через бортовой ретрансляционный комплекс (ретранслятор) космического аппарата, элементами которой являются отдельные экземпляры МРИУС распределенных локаций.

В рамках выполняемых на основе данной системы экспериментальных исследований пользователю с помощью макетов оборудования передающей земной станции (локация 1) предоставляется возможность формирования ВЧ-сигнала с заданными параметрами (число сигналов (несущих) в полосе, скорость передачи данных, используемый вид модуляции и др.), например, фрагмент комплексной огибающей которого оцифровывается и передается по вычислительной сети в локацию 2. В локации 2 данный сигнал с использованием соответствующих технических средств воспроизводится и подается на вход макета ретранслятора или имитатора радиоканала, имитирующего условия прохождения сигнала в космической системе с учетом неравномерности амплитудно-частотной характеристики и группового времени запаздывания ретранслятора, нелинейности амплитудной характеристики его усилителя, фазовых шумов гетеродинов и др. Оцифрованный фрагмент сигнала с выхода ретранслятора передается по вычислительной сети в локацию 3. С использованием технических средств в локации 3 на основе полученного цифрового «слепок» сигнал может быть сформирован на входе макетов оборудования приемной станции для оценки возможности его обнаружения, приема, обработки и др., либо оценки обобщенных качественных характеристик исследуемой системы связи (вероятность битовой ошибки, модуль вектора ошибки и пр.) [25].

При этом также следует отметить, что при описанном подходе макеты оборудования, расположенные в локациях 1–3, продолжают являться объектами полноценных экспериментальных исследований для обучаемых каждой локации, проводимых в рамках соответствующих учебных курсов.

Рассмотрим систематизированные области экспериментальных исследований, приведенные в *таблице 1*. Реализация исследований в соответствии с пп. 2.1–2.2, 2.6 основывается на создании образца МРИУС в виде открытой платформы (с открыты-

ми программными интерфейсами). На основе этой платформы обучаемый может апробировать и провести функциональное тестирование собственно-ручно разработанного и созданного графического интерфейса пользователя для удаленного управления АЛМ по различным каналам связи на основе ПЭВМ, смартфона, ноутбука [26, 27] (с использованием различных программных технологий и протоколов связи, на базе различных операционных систем), а также экспериментально апробировать разработанные алгоритмы диспетчеризации.

Исследование статистических закономерностей и динамики функционирования МРИУС как системы массового обслуживания (пп. 2.3–2.5) реализуется на базе обработки массива соответствующих статистических данных. К таким данным относятся моменты времени отправки заданий и получения результатов их выполнения, момент времени поступления, начала и завершения выполнения задания на измерительно-управляющей ЭВМ, длительности выполняемых функциональных операций и другие. Такие данные могут быть получены в процессе функционирования образцов МРИУС, например, на основе разработанного и апробированного специализированного программного модуля, обеспечивающего измерение и их централизованные накопление и хранение [14].

Таким образом, информативность МРИУС как ТСО и их дальнейшее мультidisциплинарное применение в рамках обозначенных направлений должны закладываться на этапе проектирования посредством их реализации на основе пула современных технологий, а также с использованием научного подхода к их созданию с применением современных средств автоматизированного проектирования. Такой подход должен обеспечивать гармонизацию решаемых функциональных пользовательских задач (в соответствующей предметной области АЛМ и планируемых направлений междисциплинарного применения) и требуемых технических характеристик функционирования как СМО (среднее время реакции на запросы пользователей, число одновременно работающих пользователей и др.).

2. Методология проектирования МРИУС как специализированных IoT-систем

На основе результатов проведенных исследований была разработана предметно-ориентированная методология проектирования МРИУС, объединяющая как известные [28], так и новые разработанные подходы:

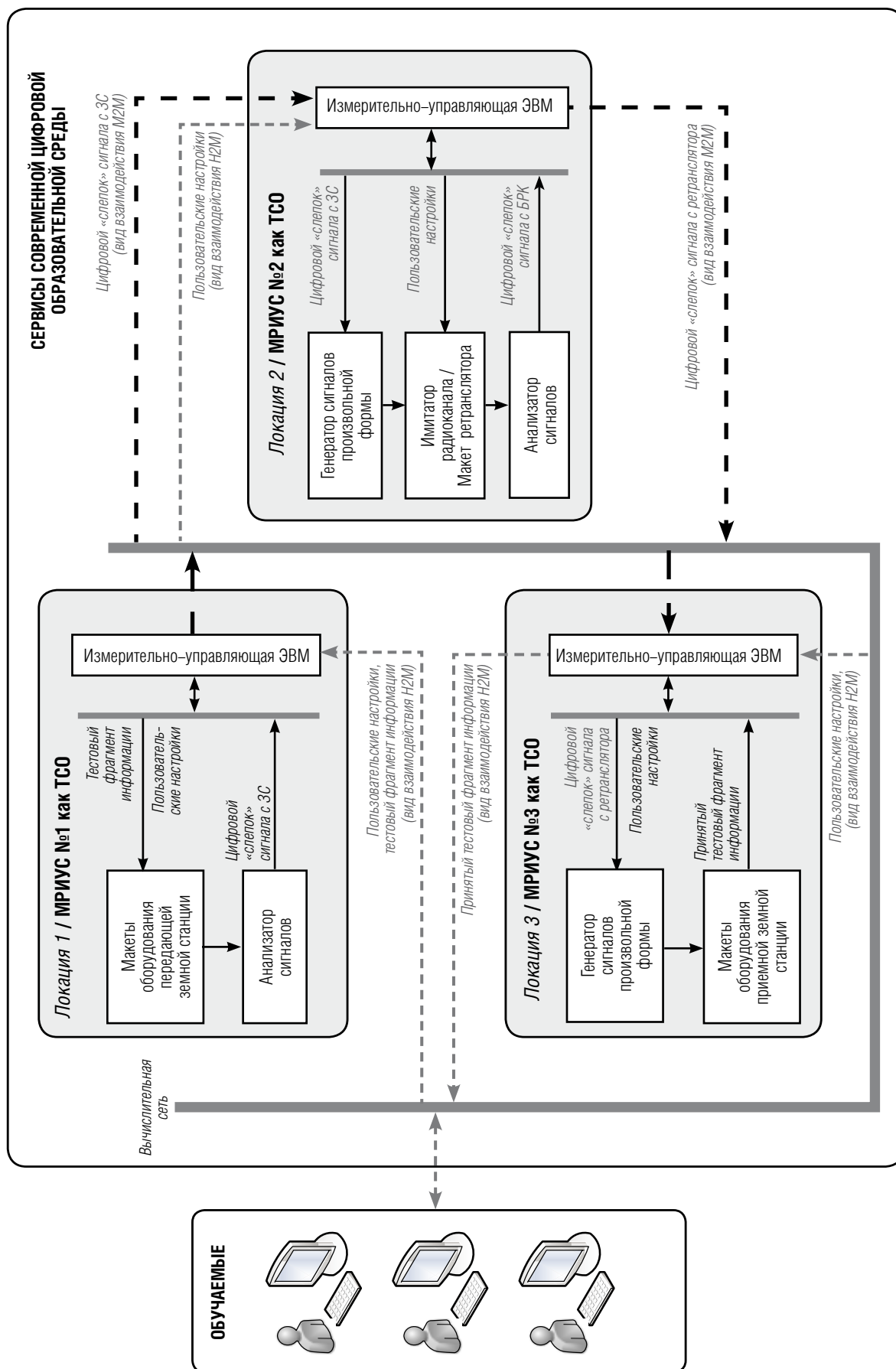


Рис. 2. Пример построения исследовательского макета системы спутниковой связи в виде IoT-суперсистемы, элементами которой являются отдельные экземпляры МРИУС распределенных локаций

- ◆ методы повышения оперативности функционирования (метод оптимизации операций управления, метод временного разделения многократных измерений, метод распараллеливания функциональных операций);
- ◆ математические модели динамики функционирования, учитывающие гетерогенность поведения и обслуживания пользователей, вариативность реализуемых алгоритмов диспетчеризации, зависимость длительностей выполняемых функциональных операций измерения и управления от параметрического и функционального содержания заданий и др., а также реализующие их адаптируемые программные модули;
- ◆ методики решения частных задач (проектирование аппаратно-программного обеспечения, формирование граничных параметров сценария диалога, расчета функций параметрической чувствительности и др.), которые обеспечивают решения проектных задач в формализованном виде;
- ◆ комплекс эвристических решений (набор базовых структур аппаратного построения МРИУС, алгоритмы диспетчеризации разделяемого ресурса, алгоритмы спектрального измерения параметров сигналов и реализующие их программные модули, базовые унифицированные виртуальные приборы, базовые шаблоны проектирования графического интерфейса пользователя).

Обобщенная структурная схема методологии приведена на *рисунке 3*. На основе выстроенного взаимодействия компонентов методологии и комплексного применения современных средств автоматизированного проектирования в рамках методики проектирования МРИУС как специализированной IoT-системы осуществляется поиск проектного решения с улучшенными показателями технического уровня. Также в целях информационной поддержки процесса проектирования предложен информационный состав базы знаний инженерных решений, аккумулирующей и реализующей возможность применение опыта предыдущих разработок.

Результаты проектирования в виде комплектов электронной программной и конструкторской документации (ПД и КД), результаты анализа различных характеристик МРИУС и др. интегрируются в составе интерактивного электронного технического руководства (ИЭТР) [29]. ИЭТР для соответствующих образцов МРИУС также являются вспомогательным элементом учебного процесса в рамках внедрения практик в области технологии сквозного автома-

тизированного проектирования. Сопровождение процесса опытной/промышленной эксплуатации промышленных образцов МРИУС осуществляется на основе предложенной методики адаптации к изменяющимся условиям эксплуатации (увеличение количества пользователей, вариативность «времени размышления» пользователей, развертывание новых локаций доступа и др.) [14].

Практическое применение предложенной методологии и методики адаптации МРИУС способствовало сокращению сроков их проектирования на 20–30%, позволило в 3–6 раз расширить множество альтернативных вариантов проектных решений, синтезируемых в соответствии с заданными требованиями технического задания, увеличить в ряде случаев число одновременно обслуживаемых терминалов пользователей на 30% и более, а также создать ряд промышленных образцов МРИУС в целях ресурсного обеспечения процессов натурных экспериментальных исследований как ОСПК, так и производственной деятельности. Предложенная методология в комплексе с методикой адаптации позволяет обеспечить эффективное применение МРИУС в составе сервиса “laboratory as a service” с одновременной минимизацией затрат на аппаратно-программное обеспечение за счет реализации концепции мультиарендности. Высокая эффективность такого сервиса, реализуемого на базе МРИУС, способствует широким возможностям применения методов электронной педагогики при подготовке и переподготовке инженерных кадров для цифровой экономики.

Сервис “laboratory as a service” позволяет перейти нескольким учебным заведениям или учебным центрам (как эксплуатирующим организациям) от развертывания и обслуживания высокотехнологичного экспериментального оборудования к его временной аренде и удаленному доступу через интернет (реализации экономической модели совместного потребления – sharing economy) [12–15, 30].

Заключение

Рассмотренные направления мультидисциплинарных исследований и предложенный подход к их реализации на базе МРИУС обеспечиваются за счет следующих факторов:

- ◆ широкий спектр экспериментальных лабораторных исследований в составе современных образовательных сред, позволяющий обучающимся изучать различные аспекты IoT-систем (принци-

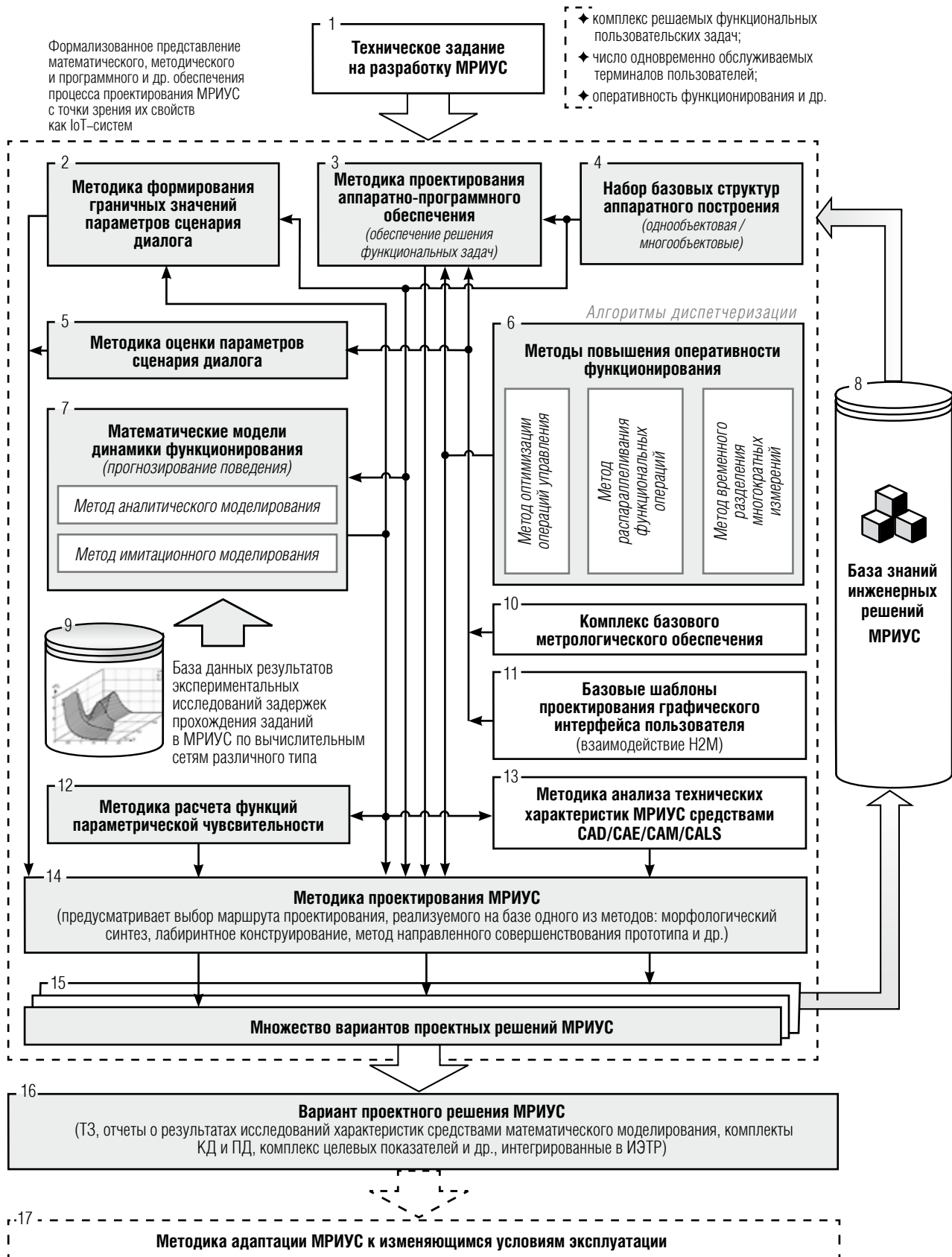


Рис. 3. Методология проектирования МРИУС с заданными техническими характеристиками:
блоки 2–7, 9–11 – решение проектных задач в областях исследований по пп. 2.1–2.8 таблицы 1;
блоки 3, 8, 11, 13, 14, 16, 17 – решение проектных задач в областях исследований по пп. 3.1, 3.2 таблицы 1

пы функционирования, методы проектирования, целевое назначение) с учетом ключевых аспектов стратегии «Образования 4.0» и направлений развития электронного приборостроения;

- ♦ развитие у обучающихся в процессе учебного и научного экспериментов практик, ориентированных на решение широкого круга задач в области цифровизации различных сфер человеческой деятельности;
- ♦ развитие у обучающихся в рамках профессиональных компетенций таких навыков, как «системные способности», «когнитивные способности» и «решение сложных проблем», которые

востребованы в основных секторах цифровой экономики;

- ♦ разработка образцов МРИУС с заданными характеристиками (число обслуживаемых пользователей, время реакции) за счет взаимосвязанных компонентов предметно-ориентированной методологии, которые включают предложенные методы, модели, методики и специализированное программное обеспечение, а также обеспечивают эффективное применение современных систем автоматизированного проектирования;
- ♦ повышение эффективности и функциональных возможностей сервиса “laboratory as a service”

для процессов подготовки и переподготовки кадров в области электронного приборостроения на базе методов электронной педагогики. ■

Литература

1. Кузьминов Я.И. Вирусная революция: как пандемия изменит наш мир. [Электронный ресурс]: <https://www.rbc.ru/opinions/society/27/03/2020/5e7cd7799a79471ed230b774> (дата обращения: 29.03.2020).
2. Cyber-physical laboratories in engineering and science education / M.E. Auer, A.K.M. Azad, A. Edwards, T. De Jong (eds.). Springer, 2018. DOI: 10.1007/978-3-319-76935-6.
3. IOT enabling measurement applications in Industry 4.0: Platform for remote programming ATES / L. Angrisani [et al.] // Workshop on Metrology for Industry 4.0 and IoT. Brescia, Italy, 16–18 April 2018. P. 40–45. DOI: 10.1109/METROI4.2018.8428326.
4. Elsaadany A., Soliman M. Experimental evaluation of Internet of Things in the educational environment // International Journal of Engineering Pedagogy. 2017. Vol. 7. No 3. P. 50–60. DOI: 10.3991/ijep.v7i3.7187.
5. Chen H. Applications of cyber-physical system: A literature review // Journal of Industrial Integration and Management. 2017. Vol. 2. No 3, Article no 1750012. DOI: 10.1142/S2424862217500129.
6. Lee E.A. The past, present and future of cyber-physical systems: A focus on models // Sensors. 2015. Vol. 15. No 3. P. 4837–4869. DOI: 10.3390/s150304837.
7. SerdarAsan S., Isikli E. Engineering education trends in the digital era. IGI Global, 2020.
8. Заславская О.Ю., Кириллов А.И. Новые возможности информатизации образования – «Интернет вещей» // Вестник Российского университета дружбы народов. Серия «Информатизация образования». 2017. Т. 14. № 2. С. 140–147. DOI: 10.22363/2312-8631-2017-14-2-140-147.
9. Heradio R., de la Torre L., Dormido S. Virtual and remote labs in control education: A survey // Annual Reviews in Control. 2016. Vol. 42. P. 1–10. DOI: 10.1016/j.arcontrol.2016.08.001.
10. Gomes L., Bogosyan S. Current trends in remote laboratories // IEEE Transactions on Industrial Electronics. 2009. Vol. 56. No 12. P. 4744–4756. DOI: 10.1109/TIE.2009.2033293.
11. Hussin A.A. Education 4.0 made simple: ideas for teaching // International Journal of Education & Literacy Studies. 2018. Vol. 6. No 3. P. 92–98. DOI: 10.7575/aiac.ijels.v6n.3p.92.
12. Laboratory as a Service (LaaS): A model for developing and implementing remote laboratories as modular components / M. Tawfik [et al.] // 11th Conference on Remote Engineering and Virtual Instrumentation (REV 2014). Porto, Portugal, 26–28 February 2014. P. 11–20. DOI: 10.1109/REV.2014.6784238.
13. Дешко И.П., Кряженков К.Г. Лабораторный практикум как сервис в курсах по информационно-коммуникационным технологиям // Перспективы науки и образования. 2015. № 1 (13). С. 70–74.
14. Комаров В.А., Сарафанов А.В., Тумковский С.Р. Многопользовательские распределенные измерительно-управляющие системы как элемент современной цифровой образовательной среды // Информационно-управляющие системы. 2019. № 2. С. 83–94. DOI: 10.31799/1684-8853-2019-2-83-94.
15. Комаров В.А., Сарафанов А.В., Тумковский С.Р. Сравнение эффективности методов повышения пропускной способности экспериментального оборудования с удаленным доступом // Информационно-управляющие системы. 2019. № 6. С. 68–76. DOI: 10.31799/1684-8853-2019-6-68-76.
16. ГОСТ Р 53626-2009. Информационно-коммуникационные технологии в образовании. Технические средства обучения. Общие положения. М.: Стандартинформ, 2010.
17. Ray P.P. A survey on Internet of Things architectures / Journal of King Saud University – Computer and Information Sciences. 2018. Vol. 30. No 3. P. 291–319. DOI: 10.1016/j.jksuci.2016.10.003.
18. El-Hasan T.S. Internet of Thing (IoT) based remote labs in engineering // 2019 6th International Conference on Control, Decision and

- Information Technologies (CoDIT). Paris, France, 23–26 April 2019. P. 976–982. DOI: 10.1109/CoDIT.2019.8820591.
19. Remote control of laboratory equipment for basic electronics courses: A LabVIEW-based implementation / B. Popovic [et al.] // Computer Applications in Engineering Education. 2013. Vol. 21. No E1. P. E110–E120. DOI: 10.1002/cae.20531.
 20. ISO/IEC 20924:2018. Information technology – Internet of Thing (IoT) – Vocabulary. [Электронный ресурс]: <https://www.iso.org/standard/69470.html> (дата обращения: 15.10.2020).
 21. ISO/IEC 30141:2018. Internet of Thing (IoT) – Reference Architecture. [Электронный ресурс]: <https://www.iso.org/standard/65695.html> (дата обращения: 15.10.2020).
 22. Kucuk K. RTWiFi-Lab: A real-time Wi-Fi laboratory platform on USRP and LabVIEW for wireless communications education and research // Computer Applications in Engineering Education. 2018. Vol. 26. No 1. P. 111–124. DOI: 10.1002/cae.21865.
 23. Bykhovsky D. Teaching wireless channel modeling with software defined radio // Computer Applications in Engineering Education. 2020. Vol. 28. No 2. P. 314–323. DOI: 10.1002/cae.22195.
 24. Toyoda Y., Koike N., Li Y. An FPGA-based remote laboratory: Implementing semi-automatic experiments in the hybrid cloud // 2016 13th International Conference on Remote Engineering and Virtual Instrumentation (REV). Madrid, Spain, 24–26 February 2016. P. 24–29. DOI: 10.1109/REV.2016.7444435.
 25. Скляр Б. Цифровая связь. Теоретические основы и практическое применение. М.: Вильямс, 2003.
 26. Odeh S. A web-based remote lab platform with reusability for electronic experiments in engineering education // International Journal of Online and Biomedical Engineering. 2014. Vol. 10. No 4. P. 40–45. DOI: 10.3991/ijoe.v10i4.3729.
 27. Reconfigurable web-interface remote lab for instrumentation and electronic learning / J.M. Sierra-Fernandez [et al.] // 2020 IEEE Global Engineering Education Conference (EDUCON). Porto, Portugal, 27–30 April 2020. P. 713–717. DOI: 10.1109/EDUCON45650.2020.9125380.
 28. Кирсанов А.Ю., Салахова А.Ш. Экспериментальное исследование статистической динамики системы дистанционного управления экспериментом в многопользовательском режиме // Нелинейный мир. 2009. Т. 7. № 5. С. 405–409.
 29. ГОСТ Р 54088–2017. Интегрированная логистическая поддержка. Эксплуатационная и ремонтная документация в форме и интерактивных электронных технических руководств. Основные положения и общие требования. М.: Стандартинформ, 2018.
 30. Spreading remote laboratory scope through a federation of nodes: VISIR case / F. Garcia-Loro [et al.] // IEEE Revista Iberoamericana de Tecnologías del Aprendizaje. 2019. Vol. 14. No 4. P. 107–116. DOI: 10.1109/RITA.2019.2950131.

Об авторах

Комаров Владимир Александрович

кандидат технических наук, доцент;

начальник группы комплексного моделирования космических систем, управление проектирования космических систем и комплексов связи, ретрансляции информации и специального назначения, Акционерное общество «Информационные спутниковые системы» имени академика М.Ф. Решетнева», 662972, г. Железногорск, ул. Ленина, д. 52;

доцент базовой кафедры «Радиоэлектронная техника информационных систем», Сибирский федеральный университет, 660041, г. Красноярск, пр. Свободный, д. 79;

E-mail: VKomarov@iss-reshetnev.ru

ORCID: 0000-0001-9210-9908

Сарафанов Альберт Викторович

доктор технических наук, профессор;

директор по развитию бизнеса, департамент систем управления предприятиями, ООО «Витте Консалтинг» (группа компаний «Ай-Тек»), 117036, Москва, ул. Кедрова, д. 15;

E-mail: Sarafanov@i-teco.ru

ORCID: 0000-0003-4264-9388

IoT systems in the process of multidisciplinary training of personnel for the digital economy and their design

Vladimir A. Komarov^{a,b}

E-mail: VKomarov@iss-reshetnev.ru

Albert V. Sarafanov^c

E-mail: Sarafanov@i-teco.ru

^a JSC “Academician M.F. Reshetnev Information Satellite Systems”

Address: 52, Lenin Street, Zheleznogorsk 662972, Russia

^b Siberian Federal University

Address: 79, Svobodny Prospect, Krasnoyarsk 660041, Russia

^c Vitte Consulting Co. (I-Teco Group)

Address: 15, Kedrova Street, Moscow 117036, Russia

Abstract

In the context of digitalization of the most knowledge-intensive sectors of the domestic economy, the development of an industrial training system in the field of electronic instrumentation is of great importance. The key areas of its development with the use of information and communication technologies include the development and improvement of the technological basis for training and retraining of personnel in engineering educational programs. One of the elements of this basis is the service of multi-user remote access via the internet to a high-tech experimental equipment laboratory as a service based on internet of things (IoT) systems. Within the framework of this service, an urgent problem is to increase the functional saturation of automated stands/installations, which is currently characterized by a paucity of scientific research. The purpose of the research is to expand the areas of experimental research carried out in the mode of multiuser remote access based on specialized IoT systems. As a result, a method of multidisciplinary application of specialized IoT systems was developed. This consists of the technical implementation of possibilities for additional research: research into technologies underlying both multi-user distributed measuring and control systems and IoT systems in general; research into technologies used in their end-to-end computer-aided design; research into joint interaction of several geographically distributed automated stands/installations, implemented on the basis of a four level IoT reference architecture. A methodology for the design of multi-user distributed measuring and control systems as specialized IoT systems has also been developed, focused on solving multidisciplinary research problems in an interactive dialogue mode based on single sample of experimental equipment. The methodology mobilizes organizational, technical and methodological support for the process of creating such systems with specified target characteristics. In general, the method and methodology developed open up opportunities for systematically implementing the basic principles of the “Education 4.0” concept in the preparation and retraining of engineering personnel in the field of electronic instrumentation.

Key words: internet of things (IoT); IoT system; laboratory as a service; technical teaching tools; multiuser distributed measuring and control systems; design methodology.

Citation: Komarov V.A., Sarafanov A.V. (2021) IoT systems in the process of multidisciplinary training of personnel for the digital economy and their design. *Business Informatics*, vol. 15, no 2, pp. 47–59.
DOI: 10.17323/2587-814X.2021.2.47.59

References

1. Kuzminov Ya.I. (2020) *The Viral Revolution: how the pandemic will change our world*. Available at: <https://www.rbc.ru/opinions/society/27/03/2020/5e7cd7799a79471ed230b774> (accessed 29 March 2020) (in Russian).
2. Auer M.E., Azad A.K.M., Edwards A., De Jong T., eds. (2018) *Cyber-physical laboratories in engineering and science education*. Springer. DOI: 10.1007/978-3-319-76935-6.
3. Angrisani L., Cesaro U., D’Arco M., Grillo D., Tocchi A. (2018) IOT enabling measurement applications in Industry 4.0: Platform for remote programming ATES. Proceedings of the *Workshop on Metrology for Industry 4.0 and IoT, Brescia, Italy, 16–18 April 2018*, pp. 40–45. DOI: 10.1109/METROI4.2018.8428326.
4. Elsaadany A., Soliman M. (2017) Experimental evaluation of Internet of Things in the educational environment. *International Journal of Engineering Pedagogy*, vol. 7, no 3, pp. 50–60. DOI: 10.3991/ijep.v7i3.7187.
5. Chen H. (2017) Applications of cyber-physical system: A literature review. *Journal of Industrial Integration and Management*, vol. 2, no 3, article no 1750012. DOI: 10.1142/S2424862217500129.
6. Lee E.A. (2015) The past, present and future of cyber-physical systems: A focus on models. *Sensors*, vol. 15, no 3, pp. 4837–4869. DOI: 10.3390/s150304837.
7. SerdarAsan S., Isikli E. (2020) *Engineering education trends in the digital era*. IGI Global.
8. Zaslavskaya O.Y., Kirillov A.I. (2017) New features of education informatization – “internet of things”. *RUDN Journal of Informatization in Education*, vol. 14, no 2, pp. 140–147 (in Russian). DOI: 10.22363/2312-8631-2017-14-2-140-147.
9. Heradio R., de la Torre L., Dormido S. (2016) Virtual and remote labs in control education: A survey. *Annual Reviews in Control*, vol. 42, pp. 1–10. DOI: 10.1016/j.arcontrol.2016.08.001.
10. Gomes L., Bogosyan S. (2009) Current trends in remote laboratories. *IEEE Transactions on Industrial Electronics*, vol. 56, no 12, pp. 4744–4756. DOI: 10.1109/TIE.2009.2033293.

11. Hussin A.A. (2018) Education 4.0 made simple: ideas for teaching. *International Journal of Education & Literacy Studies*, vol. 6, no 3, pp. 92–98. DOI: 10.7575/aiac.ijels.v6n.3p.92.
12. Tawfik M., Salzmann C., Gillet D., Lowe D., Saliah-Hassane H., Elio Sancristobal E., Castro M. (2014) Laboratory as a Service (LaaS): A model for developing and implementing remote laboratories as modular components. Proceedings of the *11th International Conference on Remote Engineering and Virtual Instrumentation (REV 2014)*, Porto, Portugal, 26–28 February 2014, pp. 11–20. DOI: 10.1109/REV.2014.6784238.
13. Deshko I.P., Kriazhenkov K.G. (2015) Labs as a service at study ICT courses. *Perspectives of Science and Education*, no 1 (13), pp. 70–74 (in Russian).
14. Komarov V.A., Sarafanov A.V., Tumkovskiy S.R. (2019) Multi-user distributed information-control systems as an element of modern digital educational environment. *Information and Control Systems*, no 2, pp. 89–94 (in Russian). DOI: 10.31799/1684-8853-2019-2-83-94.
15. Komarov V.A., Sarafanov A.V., Tumkovskiy S.R. (2019) Comparison of the effectiveness of methods to increase the throughput of experimental equipment with remote access. *Information and Control Systems*, no 6, pp. 68–76 (in Russian). DOI: 10.31799/1684-8853-2019-6-68-76.
16. Standardinform (2010) *GOST R 53626-2009. Information and communication technologies in education. Training equipment. General statements*. Moscow: Standardinform (in Russian).
17. Ray P.P. (2018) A survey on Internet of Things architectures. *Journal of King Saud University – Computer and Information Sciences*, vol. 30, no 3, pp. 291–319. DOI: 10.1016/j.jksuci.2016.10.003.
18. El-Hasan T.S. (2019) Internet of Thing (IoT) based remote labs in engineering. Proceedings of the *2019 6th International Conference on Control, Decision and Information Technologies (CoDIT)*, Paris, France, 23–26 April 2019, pp. 976–982. DOI: 10.1109/CoDIT.2019.8820591.
19. Popovic B., Popovic N., Mijic D., Stankovski S., Ostojic G. (2013) Remote control of laboratory equipment for basic electronics courses: A LabVIEW-based implementation. *Computer Applications in Engineering Education*, vol. 21, no E1, pp. E110–E120. DOI: 10.1002/cae.20531.
20. ISO (2018) *ISO/IEC 20924:2018. Information technology – Internet of Thing (IoT) – Vocabulary*. Available at: <https://www.iso.org/standard/69470.html> (accessed 15 October 2020).
21. ISO (2018) *ISO/IEC 30141:2018. Internet of Thing (IoT) – Reference Architecture*. Available at: <https://www.iso.org/standard/65695.html> (accessed 15 October 2020).
22. Kucuk K. (2018) RTWiFi-Lab: A real-time Wi-Fi laboratory platform on USRP and LabVIEW for wireless communications education and research. *Computer Applications in Engineering Education*, vol. 26, no 1, pp. 111–124. DOI: 10.1002/cae.21865.
23. Bykhovsky D. (2020) Teaching wireless channel modeling with software defined radio. *Computer Applications in Engineering Education*, vol. 28, no 2, pp. 314–323. DOI: 10.1002/cae.22195.
24. Toyoda Y., Koike N., Li Y. (2016) An FPGA-based remote laboratory: Implementing semi-automatic experiments in the hybrid cloud. Proceedings of the *2016 13th International Conference on Remote Engineering and Virtual Instrumentation (REV)*, Madrid, Spain, 24–26 February 2016, pp. 24–29. DOI: 10.1109/REV.2016.7444435.
25. Sklar B. (2003) *Digital communication. Theoretical bases and practical application*. Moscow: Williams (in Russian).
26. Odeh S. (2014) A web-based remote lab platform with reusability for electronic experiments in engineering education. *International Journal of Online and Biomedical Engineering*, vol. 10, no 4, pp. 40–45. DOI: 10.3991/ijoe.v10i4.3729.
27. Sierra-Fernandez J.M., Florencias-Oliveros O., Espinosa-Gavira M.J., Palomares-Salas J.C., Agüera-Perez A., Gonzalez-de-la-Rosa J.J. (2020) Reconfigurable web-interface remote lab for instrumentation and electronic learning. Proceedings of the *2020 IEEE Global Engineering Education Conference (EDUCON)*, Porto, Portugal, 27–30 April 2020, pp. 713–717. DOI: 10.1109/EDUCON45650.2020.9125380.
28. Kirsanov A.Yu., Salakhova A.Sh. (2009) Experimental study of the statistical dynamics of a remote control system for an experiment in multiuser mode. *Non-Linear World*, vol. 7, no 5, pp. 405–409 (in Russian).
29. Standatdinform (2018) *GOST R 54088-2017. Integrated logistic support. Operating and maintenance documentation in interactive electronic technical manuals format. General provisions and general requirements*. Moscow: Standatdinform (in Russian).
30. Garcia-Loro F., Baizan P., Blazquez-Merino M., Plaza P., Aroca A.M., Orduna P., Cristobal E.S., Castro M. (2019) Spreading remote laboratory scope through a federation of nodes: VISIR case. *IEEE Revista Iberoamericana de Tecnologías del Aprendizaje*, vol. 14, no 4, pp. 107–116. DOI: 10.1109/RITA.2019.2950131.

About the authors

Vladimir A. Komarov

Cand. Sci. (Tech.), Associate Professor;

Head of the Complex Modeling Space Systems Group, Department of Design of Space Systems and Communication Complexes, Information Relaying and Special Purpose, JSC “Academician M.F. Reshetnev Information Satellite Systems”, 52, Lenin Street, Zheleznogorsk 662972, Russia;

Associate Professor, Basic Department of Radio-Electronic Equipment of Information Systems, Siberian Federal University, 79, Svobodny Prospect, Krasnoyarsk 660041, Russia;

E-mail: VKomarov@iss-reshetnev.ru

ORCID: 0000-0001-9210-9908

Albert V. Sarafanov

Dr. Sci. (Tech.), Professor;

Business Development Director, Department of Enterprise Management Systems, Vitte Consulting Co. (I-Teco Group), 15, Kedrova Street, Moscow 117036, Russia;

E-mail: Sarafanov@i-teco.ru

ORCID: 0000-0003-4264-9388

Проблема потери решений в задаче поиска схожих документов: Применение терминологии при построении векторной модели корпуса

Ф.В. Краснов 

E-mail: fkrasnov@naumen.ru

И.С. Смазневич 

E-mail: ismaznevich@naumen.ru

Е.Н. Баскакова 

E-mail: enbaskakova@naumen.ru

NAUMEN R&D

Адрес: 620028, г. Екатеринбург, ул. Татищева, д. 49А

Аннотация

В статье рассматривается задача поиска схожих по смыслу текстовых документов в корпусе. Исследуется проблема невыявления алгоритмом TF-IDF части решений, возникающая при разработке прикладных интеллектуальных информационных систем: потеря пар, схожих согласно человеческой оценке, но получающих низкую оценку схожести от программы. Предложена модификация алгоритма с заменой общего словаря на словарь специализированных терминов. Добавление тезаурусов при построении векторной модели корпуса, основанной на ранжирующей функции, не было ранее исследовано; применение тезаурусов до сих пор изучалось лишь для улучшения тематической модели. Цель работы — повысить качество решения, минимизируя потерю значимой его части и не добавляя «ложно-схожие» пары документов, за счет применения при векторном разложении TF-IDF словаря терминов, выделенного из текста анализируемых документов. Эксперимент проведен поочередно на двух корпусах структурированных нормативно-технических документов, объединенных тематически: стандартов в отношении информационных технологий и в сфере железных дорог. Словарь терминов составлен при автоматическом анализе текста рассматриваемых документов методами выделения именованных сущностей, основанных на правилах. Продемонстрировано, что разложение TF-IDF по словарю терминов дает больше релевантных результатов для исследуемой задачи, что подтвердило выдвинутую гипотезу. Предложенный метод в меньшей степени зависит от недостатков текстового слоя (таких как ошибки распознавания), чем расчет близости документов по полному словарю корпуса. Определены факторы, способные повлиять на качество решения: способ составления словаря терминов, выбор диапазона n -грамм для словаря, корректность формулировки терминов и обоснованность их включения в глоссарий документа. Полученные выводы могут использоваться при решении прикладных задач, связанных с поиском близких по смыслу документов, таких как семантический поиск с учетом предметной области, корпоративный поиск в многопользовательском режиме, обнаружение скрытого плагиата, выявление противоречий в коллекции документов, определение новизны в документах при построении базы знаний.

Ключевые слова: схожесть документов; семантическая близость документов; применение тезаурусов; векторная модель корпуса; прикладные интеллектуальные информационные системы; объяснимость алгоритма; оценка схожести; интеллектуальный анализ текста.

Цитирование: Краснов Ф.В., Смазневич И.С., Баскакова Е.Н. Проблема потери решений в задаче поиска схожих документов: Применение терминологии при построении векторной модели корпуса // Бизнес-информатика. 2021. Т. 15. № 2. С. 60–74. DOI: 10.17323/2587-814X.2021.2.60.74

Введение

Среди задач, которые отличают интеллектуальные прикладные информационные системы от систем автоматизации бизнес-процессов, имеет место задача обнаружения инсайтов в большом объеме информации, в частности, в базе текстовых документов компании. Эта задача включает в себя поиск «близких по смыслу» документов. Для решения данной задачи строится семантическая модель текстового корпуса, в рамках которой схожесть документов определяется как расстояние между векторами документов.

Одна из проблем заключается в возможной потере части пар документов, которые схожи с точки зрения человеческой оценки, но не удовлетворяют условию превышения порогового значения схожести, установленного в прикладной интеллектуальной информационной системе. Это приводит к задаче обнаружения той части результатов, которые не показывают достаточной степени схожести при существующих методах, но должны быть учтены с точки зрения эксперта — пользователя системы («истинно-схожих» пар документов).

При расчете векторной модели корпуса текстовых документов используются разные словари, от характеристик и ограничений которых зависит качество решения. В частности, влияние оказывают степень соответствия (сфокусированность) словаря предметной области корпуса, доля часто употребляемых и редких слов, выбор диапазона n -грамм и другие параметры.

Следует отметить, что при слишком сильном расширении словаря или чрезмерном снижении порогового значения схожести с целью включить в множество решений указанные «истинно-схожие» пары документов результат снова ухудшается. Это объясняется тем, что вместе с возвратом потерянной части решений в их число включаются и лишние, «ложно-схожие» решения, которые обнаруживают близость за счет малозначимой части словаря (слов с низким весом для семантики в рамках данного корпуса). Одним из методов, позволяющих

добиться баланса между включением в решение лишних пар документов и потерей существенной части результатов, является использование тезаурусов предметной области при построении модели текстового корпуса.

Многие прикладные задачи требуют вычисления показателей сходства между экземплярами текстов и их составных частей — параграфов, предложений. Наиболее очевидный пример — когда пользователь ищет информацию в системе и поисковая система сопоставляет текст запроса с текстами ранее сохраненных документов, чтобы выдать наиболее релевантный документ. Запрос пользователя представляет собой короткий текст, и система выводит наиболее схожие по тексту документы с помощью функции ранжирования.

Кроме модулей семантического поиска, содержащая схожесть документов используется в следующих ИТ-решениях:

- ◆ рекомендательная система для авторов, определяющая наиболее подходящий журнал для публикации;
- ◆ система маршрутизации входящих заявок, подбирающая эксперта в соответствии с ранее обработанными им документами;
- ◆ программный модуль для формирования проектных команд по определенному техническому заданию;
- ◆ компонент СЭД для определения пути согласования документа на основании его содержания.

Для построения матрицы схожести необходимо использовать векторное представление документов, которое может быть построено с помощью статистической меры TF-IDF, часто применяемой и в качестве функции ранжирования. Для вычисления меры TF-IDF учитывается весь словарь корпуса, поэтому доминирующее влияние на близость документов может оказывать общеупотребительная лексика, в то время как отраслевая специфика документов может оказаться утерянной. Возникает дополнительная задача по выявлению документов, близких за счет специфической терминологии. Модификация

метода может быть произведена за счет изменения набора термов, на основании которых определяется сходство документов. В частности, количество и точность полученных результатов могут быть улучшены путем встраивания тезауруса предметной области в алгоритм вычисления функции ранжирования.

Такая модификация алгоритма улучшает качество рекомендаций в информационных системах и ускоряет принятие решений пользователем. Это осуществляется за счет повышения объяснимости алгоритма, то есть сокращения времени, которое необходимо пользователю для понимания, почему система рекомендует ему некоторую пару документов как похожую [1]. В конечном счете это улучшение способствует росту доверия пользователя к рекомендациям системы и упрощает его аналитическую работу с текстовыми документами — поиск информации в корпоративных источниках, проверку на отсутствие дублей в базе, обнаружение пересечений и противоречий. Все это в совокупности приводит к сокращению временных затрат сотрудников на обработку больших объемов неструктурированных данных.

Целью настоящей работы является улучшение модели на основе ранжирующих функций за счет использования в качестве тезауруса словаря терминов предметной области. Авторы сфокусировались на корпусах структурированных текстовых документов в определенной предметной области. К таким документам, например, относятся нормативно-технические базы организаций [2], договорные и расходные документы, резюме кандидатов, ГОСТы и многие другие. Рассматривалась прикладная задача нахождения в текстовом корпусе схожих по смыслу документов.

Исследовалась проблема потери части решения при расчете близости документов с помощью ранжирующей функции. Данная проблема состоит в том, что некоторые пары документов являются схожими согласно человеческой оценке, однако программа не определяет их как таковые, поскольку они показывают низкую степень схожести по алгоритму TF-IDF, даже с учетом оптимизации словаря за счет удаления наиболее часто и наиболее редко встречающихся слов. На практике такая проблема возникает при разработке прикладных интеллектуальных информационных систем, где ставится задача найти в корпусе документов достаточно близкие, то есть такие, степень схожести которых превышает некоторый установленный

порог и, следовательно, является значимой для пользователя. В работе предложена модификация существующих методов вычисления близости документов с помощью ранжирующих функций, позволяющая избежать потери указанной части результатов.

Гипотеза исследования состоит в том, что при поиске в текстовом корпусе пар близких по смыслу документов с помощью ранжирующей функции TF-IDF та часть решения, которая теряется при векторном разложении по полному словарю корпуса, может быть обнаружена при построении векторной модели на основании словаря терминов, соответствующих предметной области. На *рисунке 1* показано графическое представление гипотезы.

Статья включает обзор имеющихся исследований в области применимости тезаурусов и проблематики их построения, описание методов исследования и экспериментального подтверждения исследовательской гипотезы, а также анализ полученных результатов.

1. Применение тезаурусов

Компьютерная лингвистика десятилетиями интересовалась возможностью применения тезаурусов при анализе текстов для выявления семантических связей в корпусе, поскольку созданные экспертами словари определений терминов и понятий обладают высокой смысловой точностью.

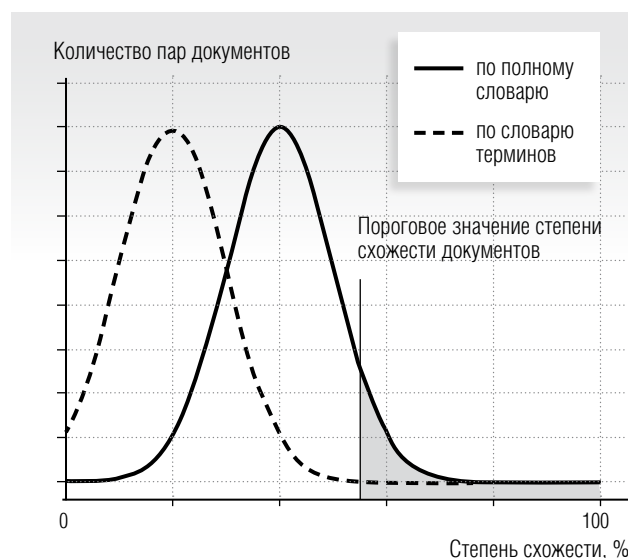


Рис. 1. Ожидаемый сдвиг графика распределения степени схожести документов при замене в алгоритме TF-IDF общего словаря на словарь терминов предметной области

Первые методы, предложенные в 70-е и 80-е годы XX века, были нацелены на создание семантических языковых моделей с помощью анализа компьютерных словарей — источников таксономических отношений «гипоним — гипероним». Для извлечения отношений из словарей использовались алгоритмы, основанные на правилах [3, 4]. Преимущество этих методов состояло в том, что семантические отношения извлекались ими из надежных источников, которые можно считать уже частично структурированными, поскольку словари работают как «имплицитные таксономии». Однако эти методы унаследовали проблемы лексикографического материала, связанные как с возможной недостоверностью данных вследствие некорректного обновления источников, так и с неполнотой словарей, поскольку даже сегодня они не всегда являются корпусными.

Во многих исследованиях проявлялся интерес к поиску со-гипонимических отношений, то есть выявлению групп слов, которые определяются с помощью одного и того же гиперонима, таких как типы систем, оборудования, связей [5, 6]. Такие слова считаются парадигматически связанными. Это означает, что они имеют тенденцию повторяться в сходных синтагматических контекстах и предположительно обладают общими семантическими признаками.

Другая стратегия выявления семантических отношений в корпусе состоит в создании языковой модели на основе статистического анализа текстов без использования предварительно полученных знаний о связях словаря. В работе [7] для выделения пар гипоним-гипероним строится направленный граф встречаемости термов, а если они слишком редкие, то в частотное распределение внедряется смещение: вводится отношение аналогии между гипонимами на основе состава и морфологических признаков термов. Например, если статистически установлено, что синдром Аспергера и синдром Карпентера (гипонимы) — заболевания (гипероним), то терм «синдром Меретойи» имеет тот же гипероним «заболевание»; к нему же может быть отнесено и слово «артрит» по аналогии с «гастритом». Авторы статьи [8] соединяют обе стратегии для создания гипернимических цепочек с использованием подходов, основанных на дистрибутивной семантике.

Потенциальная ограниченность количественных методов состоит в недостатке точности, однако проблема решается путем увеличения объема линг-

вистических данных. Поэтому такой подход считается эффективным и является довольно популярным. Кроме того, будучи независимым от языка, он может быть легко реплицирован и использован для создания многоязычных ресурсов.

В рамках статистического подхода для решения многих прикладных задач строятся векторные модели корпуса документов на основании ранжирующих функций. Часто используемая функция ранжирования TF-IDF, которая основана на двух понятиях: частота появления слова (терма) в документе (term frequency, TF) и важность этого слова для всего набора документов — обратная частота вхождения слова во все документы корпуса (inverse document frequency, IDF). Метод был предложен в 1970-х годах [9–12] и до сих пор широко применяется для анализа схожести текстов, поскольку позволяет сохранить определенные их свойства при проекции текстовых данных на числовое пространство. Измерение схожести может быть сделано через вычисление близости между TF-IDF векторами, например, по косинусной мере. Принцип сравнения с помощью TF-IDF можно назвать «пословным», так как в разреженном TF-IDF векторе число компонент соответствует числу термов в словаре. Однако набор термов, определяющих сходство пары документов, является результатом частотной дискриминации, а не содержания этих документов. Ранжирование на основе TF-IDF используется в различных приложениях, таких как кластеризация документов [13–16] и тематическое моделирование [17, 18].

Использование тезаурусов при анализе корпуса текстов методами дистрибутивной семантики рассмотрено в ряде исследований. В частности, в разное время авторы обращались к проблематике работы с тезаурусом на одном из шагов алгоритма в задаче текстового поиска. Например, в статье [19] описан подход к автоматической индексации тезисов статей с использованием тезауруса предметной области. Авторами работы [20] представлен способ автоматического извлечения ключевых фраз с помощью семантической информации о терминах и фразах, собранных из предметно-ориентированного тезауруса. В работе [21] применяется политематический тезаурус для целей семантического сопоставления понятийных образов.

Заметное число исследований посвящено улучшению параметров тематических векторных моделей корпуса, в которых векторное разложение документов сделано по набору тем, выделенных в

результате анализа текста на всем корпусе. Например, в работе [22] улучшение тематической модели производится с помощью искусственного увеличения совстречаемости синонимов, а для усиления когерентности тем модели в исследовании [23] информация о синонимах вносится в априорное распределение Дирихле. В работе [24] предложено такое понятие как Thesaurus-Based Topic Model и произведено сравнение различных тематических моделей. Все вышеперечисленные исследования объединяет общий результат экспериментов: исключение гипонимов и увеличение частот фраз улучшает человеческую оценку качества получаемых тем, поскольку придает больший вес более конкретным словам и фразам, содержание которых лучше определено.

В то же время возможность применения тезаурусов для улучшения семантической модели корпуса текстов, в которой матрица близости строится на основании ранжирующей функции (такой как TF-IDF), до сих пор не была достаточно изучена.

2. Построение тезаурусов

Согласно определению Международной организации по стандартизации (ISO)¹, тезаурус является словарем, формально организованным для того, чтобы установить явные априорные отношения между понятиями. Элементами тезауруса являются лексические единицы и семантические отношения (связи) между ними. Тезаурусные отношения (род – вид, часть – целое, комплекс – элемент, причина – следствие) налагаются на структуру таксономии, то есть идентифицируются как основные таксономии предметной области.

Методика написания тезаурусов определена в стандарте «ГОСТ Р ИСО 704-2010 Терминологическая работа. Принципы и методы» [25]. Исторически тезаурусы создавались для ручного индексирования документов и при их создании не принимались во внимание вопросы, связанные с автоматической индексацией. Трудность построения тезауруса, соответствующего всему тематическому многообразию индексируемой информации, делает тезаурус самодостаточным информационным продуктом, одновременно являясь основной причиной его непопулярности в современных информационных системах.

В качестве выдающегося примера специализированных тезаурусов следует отметить российский тезаурус в области сельского хозяйства, созданный Центральной научной сельскохозяйственной библиотекой [26]. Тезаурус построен как расширение словаря-классификатора ГРНТИ (Государственного рубрикатора научно-технической информации).

Трудоемкость ручного составления тезаурусов и возникающие вследствие этого проблемы можно увидеть на примере таких нормативно-технических документов, как ГОСТы. Текст большинства из них сопровождается тезаурусом – разделом с описанием терминов и определений. Организации, отвечающие за разработку отраслевых стандартов, сталкиваются с тем, что зачастую в разных документах обнаруживаются противоречащие определения одних и тех же терминов, разные названия одного понятия или конфликты между нормативными значениями показателей. Причина этого явления состоит в том, что база ГОСТов, состоящая из сотен тысяч документов, накапливалась в течение многих десятилетий. При этом из-за ограниченных возможностей поиска положения стандартов зачастую дублировались, превращаясь со временем в противоречия из-за обновления документов без учета их тематических связей с другими.

Пример 1. Рассмотрим различные определения и названия понятия «акустический (шумозащитный) экран», имеющиеся в разных нормативно-технических документах:

1. ГОСТ Р 51943-2002. Экраны акустические для защиты от шума транспорта. Методы экспериментальной оценки эффективности (2002): «**Акустический экран, экран:** Барьер (ограниченная преграда), устанавливаемый на пути распространения шума реального источника к защищаемому от шума объекту».

2. ГОСТ 32957-2014. Дороги автомобильные общего пользования. Экраны акустические. Технические требования (2014): «**Акустический экран:** Искусственная преграда, устанавливаемая на пути распространения шума от автомобильного транспорта к защищаемому от шума объекту. Типовой акустический экран представляет собой сборную конструкцию, состоящую из следующих основных частей: фундамента (если предусмотрено проектной документацией), несущей конструкции (в частности, опорных стоек) и панелей. В качестве

¹ <http://docs.cntd.ru/document/1200129056>.

дополнительных элементов используют уплотнения, поперечные профилированные балки, крепежные детали, акустические развязки, козырьки, калитки, ворота, рамы разрывов и т.п.».

3. ГОСТ 33329-2015. Экраны акустические для железнодорожного транспорта. Технические требования (2015): «**Акустический экран**: Протяженная искусственная преграда, устанавливаемая на пути распространения шума от реального источника (железнодорожного транспорта) к защищаемому от шума объекту».

4. СП 338.1325800.2018. Защита от шума для высокоскоростных железнодорожных линий. Правила проектирования и строительства (2018): «**Шумозащитный (акустический) экран (экран); ШЭ**: Сооружения в виде вертикальных или наклонных протяженных искусственных преград различной конструкции, земляных насыпей, выемок, галерей и т.п., устанавливаемых вдоль железных дорог на пути распространения шума транспортного потока к защищаемому объекту в целях снижения шума».

5. ОДМ 218.8.011-2018. Методические рекомендации по определению характеристик и выбору шумозащитных конструкций автомобильных дорог (2018): «**Шумозащитный экран (ШЭ)**: Протяженная искусственная преграда, устанавливаемая на пути распространения шума от автомобильного транспорта к защищаемому от шума объекту, ширина (или толщина) которой много меньше ее высоты, состоящая из фундамента и закрепленного на нем шумозащитного полотна».

Чтобы избежать влияния указанных недостатков тезаурусов, составленных вручную, на результаты поиска схожих документов в прикладных информационных системах, необходимо использовать автоматические методы формирования терминологических словарей для корпуса документов.

Существующие подходы к извлечению семантических отношений на основе паттернов, подобные описанным в предыдущем параграфе, используются и в вычислительной терминологии. Например, в работе [27] описаны методы экстракции, ориентированные на выявление отношений гипоним — гипероним (частное и обобщение), мероним — холоним (часть и целое), синонимии и причинно-следственных связей, которые совокупно используются для определения терминов и их отношений в рамках подхода “data fusion pipeline”. Однако вычислительная терминология сосредоточена в основном на изучении самих паттернов семантиче-

ских отношений, их описании, интерпретации и формализации их лингвистических свойств, а также на анализе паттернов за пределами возможностей их обнаружения. Однако для рассматриваемой задачи достаточным является получение набора терминов, которые используются в исследуемом корпусе документов без учета семантических взаимоотношений этих терминов.

3. Методы

Для проверки сформулированной гипотезы был проведен эксперимент с целью сравнения результатов решения задачи поиска схожих документов в корпусе, полученных двумя методами — базовым TF-IDF и его модификацией с применением тезауруса.

В рамках исследования был рассмотрен набор из N_d структурированных текстовых документов, объединенных тематическим направлением.

В качестве тезауруса (как наиболее доступный вариант) был использован словарь T из терминов предметной области, полученный в результате автоматического анализа текстов этих документов. Для этого все документы из набора были обработаны алгоритмом извлечения именованных сущностей, основанным на правилах (rule-based NER). При этом из соответствующего раздела каждого документа, где перечислены используемые термины и их описания, были извлечены списки терминов, которые затем были объединены в словарь терминов, специфичных для данного корпуса.

В словарь включались только термины, состоящие из одного или двух слов. Слова были приведены к нормальной форме, внесена информация о морфологических признаках, исключены цифры и английские символы.

Документы были преобразованы к следующему текстовому формату: набор документов представлен в виде массива строк M , в котором одна строка соответствует одному документу. Последовательность слов при этом была сохранена.

Для получения матриц «пословной» степени схожести S_{TFIDF} текстовых документов корпуса M было использовано векторное преобразование TF-IDF в нескольких вариантах, различающихся словарем:

♦ в качестве словаря были использованы все слова из корпуса документов. Было рассмотрено не-

сколько вариантов построения словаря в зависимости от пороговой частоты использования слов в рамках всего корпуса: если слово встречалось менее определенного количества раз (параметр df_{min}), то оно исключалось из словаря;

♦ в качестве словаря был использован набор специализированных терминов T .

В результате для каждого массива было получено пять вариантов матрицы A_{TFIDF} типа «документы – словарь» с размерностью $N_D \times N_T$ (число слов в словаре). Матрица близости векторов (степени схожести документов) S_{TFIDF} вычислялась по формуле:

$$S_{TFIDF} = A_{TFIDF} \times A_{TFIDF}^T.$$

На основании матрицы A_{TFIDF} , сформированной по набору специализированных терминов, был построен граф знаний о корпусе текстовых документов G_{TFIDF} . Вершины графа (документы и термины) соединялись ребрами в том случае, если значение TF-IDF разложения для соответствующей пары «документ – термин» оказывалось выше заданного порога (вес ребра равен соответствующему значению в матрице A_{TFIDF}).

4. Результаты эксперимента

Для проведения эксперимента был выбран набор юридически значимых документов из библиотеки ГОСТ: $N_{D_{IT}} = 667$ документов в формате docx, отно-

сящихся к тематике информационных технологий. После обработки документов в соответствии с выбранной методикой был получен массив текстовых строк M_{IT} . Из текста документов было выделено $T_{IT} = 4417$ терминов.

Результаты вычислений для документов по информационным технологиям показаны на рисунке 2.

В таблице 1 приведены значения математического ожидания и среднеквадратического отклонения степени схожести пар документов M_{IT} .

Таблица 1.

Математическое ожидание и среднеквадратическое отклонение степени схожести пар документов M_{IT} при различных словарях

Условия	Математическое ожидание, %	Средне-квадратическое отклонение, %
По полному словарю при $df_{min} = 0$	2,9	2,3
По полному словарю при $df_{min} = 5$	5,8	3,9
По полному словарю при $df_{min} = 10$	7,9	4,8
По полному словарю при $df_{min} = 50$	17,4	8,0
По словарю терминов	21,9	8,0

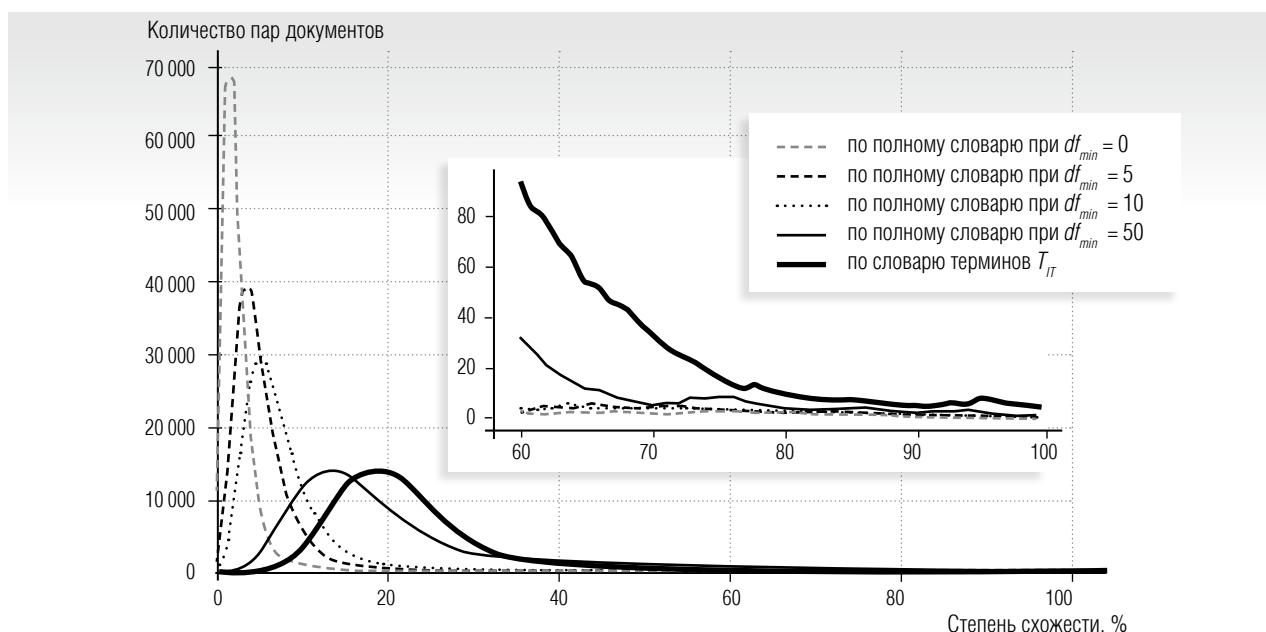


Рис. 2. Количество пар документов M_{IT} разной степени схожести при различных вариантах словаря, в т.ч. количество пар документов M_{IT} со схожестью в диапазоне 60 – 100% при разных вариантах словаря

Эксперимент подтвердил выдвинутую гипотезу: замена общего словаря на набор специализированных терминов в алгоритме на основе ранжирующей функции увеличивает количество найденных решений. На представленных графиках и из данных таблицы видно, что построение модели на специализированных терминах определяет большее число пар схожих документов, а варианты с исключением редко встречающихся в рамках корпуса слов существенных отличий не вносят. Поэтому для дальнейшего сравнения методов, включая анализ распределения близости документов, из первой группы

вариантов алгоритма был выбран один – с расчетом схожести по полному словарю при $df_{min} = 5$.

Для предметного анализа был рассмотрен диапазон схожести от 60% до 100%. Количественные характеристики полученного результата представлены в *таблице 2*.

Пример 2. В *таблице 3* приведены примеры пар документов по информационным технологиям, имеющих высокую степень близости TF-IDF по словарю терминов и низкую – по полному словарю корпуса.

Таблица 2.

Количество пар документов M_{IT} в разных интервалах степени схожести, обнаруженное алгоритмом TF-IDF с различными словарями

Степень близости документов M_{IT}	Количество пар близких документов, определенное по полному словарю при $df_{min} = 5$	Количество пар близких документов, определенное по словарю терминов
60% – 70%	42	459
70% – 80%	26	133
80% – 90%	25	55
90% – 100%	7	56

Таблица 3.

Примеры пар документов M_{IT} , имеющих высокую степень близости TF-IDF по словарю терминов и низкую – по полному словарю корпуса

№ пары	Документы	Абсолютное значение близости по TF-IDF по словарю терминов	Абсолютное значение близости по TF-IDF по полному словарю
1	ГОСТ 34.971–91 (ИСО 8822–88) Информационная технология (ИТ). Взаимосвязь открытых систем. Определение услуг уровня представления с установлением соединения	0,6944633	0,1744794
	ГОСТ Р ИСО/МЭК 9066–1–93 Системы обработки информации. Передача текста. Надежная передача. Часть 1. Модель и определение услуг		
2	ГОСТ 28147–89 Системы обработки информации. Защита криптографическая. Алгоритм криптографического преобразования	0,6806692	0,17162557
	ГОСТ Р 34.13–2015 Информационная технология (ИТ). Криптографическая защита информации. Режимы работы блочных шифров (с поправкой)		
3	ГОСТ Р 34.964–92 (ИСО 8602–87) Информационная технология (ИТ). Взаимосвязь открытых систем. Протокол транспортного уровня в режиме без установления соединения	0,7068537	0,1946876
	ГОСТ Р ИСО/МЭК 10025–3–94 Информационная технология (ИТ). Передача данных и обмен информацией между системами. Аттестационное тестирование транспортного уровня в режиме с установлением соединения при использовании услуг сетевого уровня в режиме с установлением соединения. Часть 3. Спецификация протокола административного управления тестированием		

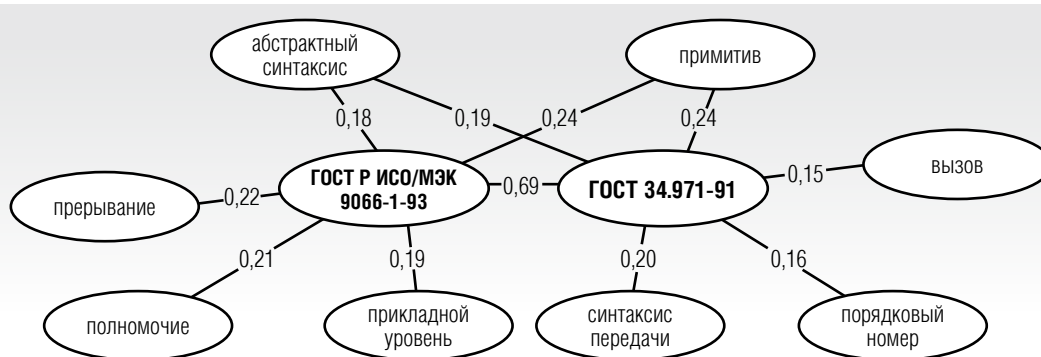


Рис. 3. Граф связей пары документов ГОСТ 34.971-91 и ГОСТ Р ИСО/МЭК 9066-1-93 и пяти наиболее значимых для каждого из них терминов

Пример 3. На рисунке 3 представлен один из вариантов графа связей документов и терминов G_{TFIDF} , а именно — его фрагмент с рассмотренными выше документами ГОСТ 34.971-91 и ГОСТ Р ИСО/МЭК 9066-1-93. Для каждого из документов показаны связи с пятью наиболее «весомыми» (значимыми) терминами.

Эксперимент был повторен на другом наборе документов: $N_{D_{RW}} = 218$ документов в формате docx, относящихся к тематике железных дорог. Для документов был сформирован массив полученных текстовых строк M_{RW} и выделен набор терминов T_{RW} .

Результаты второго эксперимента подтвердили сделанный вывод о том, что, несмотря на смещение кривой распределения числа пар по степени схожести, выявленная тенденция сохранена, а именно: при расчете близости методом TF-IDF с векторным разложением по словарю терминов в наборе обнаруживается больше схожих документов, чем при разложении по всему словарю корпуса.

Разница между матрицей схожести документов по терминам и матрицей схожести по общему словарю представлена на графиках на рисунке 4.

5. Дискуссия

Предложенный алгоритм дает существенное отличие при поиске близких документов в диапазоне схожести 60–80%. При более высокой степени схожести документов распределение близости по специализированным терминам не сильно отличается от распределения близости по полному словарю, поскольку концентрация терминов в документах приближается к концентрации общеупотребимых слов.

Из графика на рисунке 2 видно, что для данного набора документов высокая близость по словарю означает высокую близость по терминам. Однако в обратном случае поиск документов по специализированным терминам дает более высокую степень схожести.

Приведенные в таблице 3 примеры документов демонстрируют, что описанный метод расчета схожести по ранжирующей функции с разложением по словарю терминов способен снизить риск потери той части решения, где схожесть оказывается ниже

Нормированная доля пар документов, %

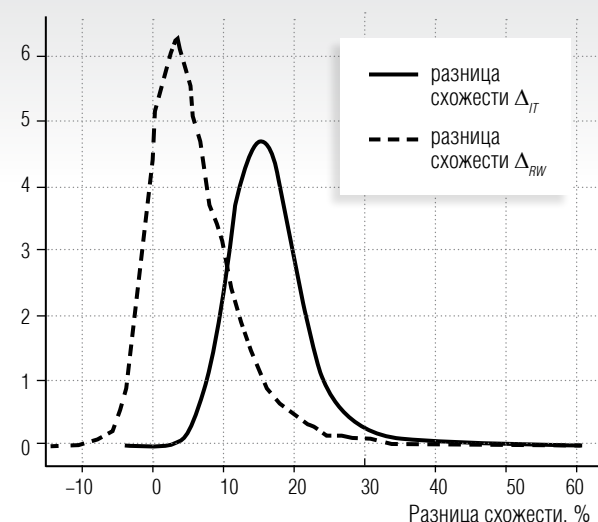


Рис. 4. Распределение разницы в степени схожести пар документов, вычисленной по двум словарям: все слова корпуса за вычетом редко встречающихся и словарь из специализированных терминов (Δ_{TT} — разница на корпусе M_{TT} , Δ_{RW} — разница на корпусе M_{RW})

установленного порога, но документы являются схожими с точки зрения пользователя системы.

Следует отметить, что в общем случае при переходе от общего словаря к терминологии при поиске схожести теоретически возможна потеря части решения – документов, схожих согласно человеческой оценке. Однако в реальных информационных системах эта потеря незначительна, так как пользователь в первую очередь сфокусирован на выявлении схожести в рамках определенной темы, то есть схожести на основании специфической для предметной области лексики. Поэтому при переходе от общего словаря к терминам возможна потеря лишь тех пар схожих документов, которые близки за счет общеупотребительной лексики. Количественные показатели таких потерь зависят от доли терминов в словаре, которая меняется от корпуса к корпусу. Для вычисления вклада общеупотребительных слов в общее значение степени схожести могут использоваться различные метрики, разработка которых выходит за рамки настоящего исследования.

В проведенном эксперименте не было обнаружено документов, у которых схожесть по общему словарю была бы больше, чем схожесть по терминам. На графике, демонстрирующем разницу между схожестью по словарю терминов и общему словарю, количество отрицательных значений незначительно, а разница между схожестью у таких пар документов составляет сотые доли процента.

Сравнение результатов двух экспериментов (на массивах M_{IT} и M_{RW}) позволило сделать дополнительные выводы и выделить факторы, влияющие на качество результата при поиске схожих документов с использованием векторного разложения по словарю терминов, в том числе созданному автоматически, на основе того же набора документов.

Неопределенность, вносимая в процесс обработки данных на шаге распознавания текста в документах, изначально полученных в графических форматах, является ключевым фактором при анализе корпуса документов. Недостаточное качество распознавания, когда часть текста теряется, снижает среднюю схожесть по TF-IDF при использовании общего словаря, однако по словарю терминов такая близость документов по-прежнему обнаруживается. Таким образом, предложенный алгоритм позволяет избежать затрат на подготовку текста для анализа – предварительную очистку текстового слоя и структурирование документа для поиска ключевых разделов.

Результат зависит и от выбора метода для автоматического составления словаря терминов. Эксперимент показал, что если термины выделяются из текста специального раздела анализируемых документов, то на результаты влияет наличие описанных терминов в содержательной части документа: если термины лишь перечислены в соответствующем разделе, но далее не используются, то значение схожести между документами искажается.

При автоматическом составлении словаря терминов схожесть также зависит от количества ошибок оформления в описании терминов в документах, поскольку в этом случае часть словаря может быть утеряна: термины используются в тексте документа, но не включены в словарь (базис для векторного разложения).

В общем случае схожесть документов по терминам зависит от выбора диапазона для n -грамм при составлении словаря терминов: длинные термины, состоящие из нескольких слов, в тексте документа с большой вероятностью ни разу не будут обнаружены целиком, когда все слова n -граммы входят в текст в нужном порядке и стоят подряд.

Пример 4. В документе из корпуса M_{RW} «ГОСТ 33798.4-2016 (IEC 60077-4:2003). Электрооборудование железнодорожного подвижного состава. Часть 4. Выключатели автоматические переменного тока. Общие технические условия» вводятся, в числе прочих, следующие термины:

♦ «Ручной рычаг включения/отключения: Рычаг для приведения автоматического выключателя во включенное или отключенное состояние ручным способом». Указанный термин ни разу не используется в содержательной части текста данного документа.

♦ «Выключатель внутренней установки: Выключатель, предназначенный только для установки и эксплуатации с защитой от неблагоприятных условий эксплуатации (ветра, дождя, снега, повышенного отложения грязи, нестандартных окружающих условий, льда и изморози)». В полном виде термин в тексте не обнаруживается, используется в другой форме только один раз: «Выключатели классифицируют: ... в соответствии с типом конструкции, то есть выключатели наружной или внутренней установки». Это означает, что данная триграмма не встречается в документе нигде, кроме самого описания термина. Таким образом, в данном случае TF (важность термина для данного документа) меняется. Поэтому при расчете схожести документов

по TF-IDF на основании словаря терминов целесообразно включать в него только униграммы и биграммы.

Наконец, если в структуре термина присутствует инверсия, то он также с высокой вероятностью не обнаруживается в тексте как исходная *n*-грамма.

Пример 5. В рассмотренном выше документе ГОСТ 33798.4-2016 определяется термин «выключатель полупроводниковый». В соответствии с нормами русского языка в связном тексте слова данного термина скорее всего поменяются местами: прилагательное разместится перед существительным. Действительно, в содержательной части документа указанный термин обнаруживается лишь один раз, без инверсии: «Для каждого выключателя указывают: вид устройства (например, воздушный выключатель, вакуумный выключатель, ..., полупроводниковый выключатель...».

Описанный эксперимент повторялся на корпусах одного типа. При проведении его на любом другом наборе документов результат может оказаться менее выразительным. Этому могут способствовать как указанные выше факторы, так и другие особенности документов: их структура, стилистика изложения, отраслевая специфика. Кроме того, в общем случае влияние на результат может оказать и тематическая сфокусированность корпуса, а также этимологические особенности терминологии, такие как использование общеупотребительной лексики в качестве узкоспециализированных терминов. Однако предварительно проведенные теоретические исследования позволяют сделать вывод о том, что при рассмотрении любого другого корпуса документов, в той или иной степени объединенных тематически, тенденция увеличения числа найденных схожих документов при разложении по словарю терминов сохранится.

Предложенный в работе подход реализован на практике — в рамках информационной системы Naumen LegalTech (продукт разрабатывается при грантовой поддержке Российского фонда развития информационных технологий). Это платформа аналитической обработки большого потока юридически значимых документов, ориентированная на пользователей, чьи интересы далеки от сферы информационных технологий и точных наук: юристов, методологов, разработчиков внутренней нормативной документации, авторов законодательных инициатив, специалистов по стандартиза-

ции, представителей отделов правовой экспертизы. Специфика работы целевой аудитории продукта такова, что в случае ошибки существует риск возникновения правовых последствий, но при этом интеллектуальные алгоритмы существенно помогают, работая в рекомендательном режиме. Для таких пользователей особенно важно доверять рекомендациям информационной системы, не погружаясь в технические детали ее реализации, а объяснимость решений является одним из ключевых факторов.

Заключение

В представленном исследовании рассмотрена возможность применения словаря специализированных терминов как частного случая тезауруса при векторном разложении корпуса документов с помощью ранжирующей функции.

Изучена проблема потери части решения в задаче поиска в корпусе текстовых документов пар элементов, близких по смыслу в значении «пословной» схожести текстов с учетом значимости термов.

В работе описан проведенный эксперимент на двух наборах нормативно-технических документов. Результаты эксперимента показали, что разложение по словарю терминов позволяет уменьшить потерю части решений, которая не удовлетворяет заданному условию превышения порогового значения степени схожести в паре документов, однако должна быть признана значимой согласно человеческой оценке.

Определены возможные факторы, влияющие на качество решения при выбранном методе поиска схожих документов в корпусе: выбор способа построения словаря терминов, ограничение длины *n*-граммы в словаре, грамматические особенности терминов.

Модели корпуса документов, основанные на специализированном терминологическом словаре, применимы в разных ИТ-решениях, учитывающих схожесть текстов: в рекомендательных системах, в модулях семантического поиска, при маршрутизации обращений. Предложенный алгоритм повышает точность и объяснимость рекомендаций в информационных системах, что ускоряет принятие решений пользователем и увеличивает эффективность его работы с большим объемом текстовых документов. ■

Литература

1. Краснов Ф.В., Смазневич И.С. Фактор объяснимости алгоритма в задачах поиска схожести текстовых документов // Вычислительные технологии. 2020. Т. 25. № 5. С. 107–123. DOI: 10.25743/ICT.2020.25.5.009.
2. Отрадных К.К., Жуков Д.О., Новикова О.А. Модель кластеризации слабоструктурированных текстовых данных // Современные информационные технологии и ИТ-образование. 2017. Т. 13. № 3. [Электронный ресурс]: <http://sitito.cs.msu.ru/index.php/SITITO/article/view/295>. DOI: 10.25559/SITITO.2017.3.439.
3. Calzolari N. An empirical approach to circularity in dictionary definitions // *Cahiers de Lexicologie*. 1977. Т. 31. №. 2. С. 118–128.
4. Amsler R.A. The structure of the Merriam-Webster pocket dictionary. Austin, TX: The University of Texas, 1980.
5. Grefenstette G. Explorations in automatic thesaurus discovery. New York: Springer Science & Business Media, 1994. DOI: 10.1007/978-1-4615-2710-7.
6. Bullinaria J.A. Semantic categorization using simple word co-occurrence statistics // ESSLLI Workshop on Distributional Lexical Semantics (ESSLLI 2008). Hamburg, Germany, 4–9 August 2008. P. 1–8.
7. Nazar R., Vivaldi J., Wanner L. Co-occurrence graphs applied to taxonomy extraction in scientific and technical corpora // *Procesamiento del Lenguaje Natural*. 2012. No 49. P. 67–74.
8. Chasing hypernyms in vector spaces with entropy / E. Santus [et al.] // 14th Conference of the European Chapter of the Association for Computational Linguistics. Gothenburg, Sweden, 26–30 April 2014. P. 38–42.
9. Jones K.S. A statistical interpretation of term specificity and its application in retrieval // *Journal of Documentation*. 1972. Vol. 28. No 1. P. 11–21.
10. Jones K.S. Index term weighting // *Information Storage and Retrieval*. 1973. Vol. 9. No 11. P. 619–633.
11. Salton G., Yang C.S. On the specification of term values in automatic indexing. Ithaca, NY: Cornell University, 1973.
12. Salton G., Wong A., Yang C.S. A vector space model for automatic indexing // *Communications of the ACM*. 1975. Vol. 18. No 11. P. 613–620.
13. Отрадных К.К., Раев В.К. Экспериментальное исследование эффективности методик векторизации текстовых документов и алгоритмов их кластеризации // Вестник Рязанского государственного радиотехнического университета. 2018. № 64. С. 73–84. DOI: 10.21667/1995-4565-2018-64-2-73-84.
14. Bafna P., Pramod D., Vaidya A. Document clustering: TF-IDF approach // 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT 2016). Chennai, India, 3–5 March 2016. P. 61–66. DOI: 10.1109/ICEEOT.2016.7754750.
15. Fomin S.A., Belousov R.L. Detecting semantic duplicates in short news items // *Business Informatics*. 2017. No 2. P. 47–56. DOI: 10.17323/1998-0663.2017.2.47.56.
16. Qaiser S., Ali R. Text mining: Use of TF-IDF to examine the relevance of words to documents // *International Journal of Computer Applications*. 2018. Vol. 181. No 1. P. 25–29. DOI: 10.5120/ijca2018917395.
17. Krasnov F., Ushmaev O. Exploration of hidden research directions in oil and gas industry via full text analysis of OnePetro digital library // *International Journal of Open Information Technologies*. 2018. Vol. 6. No 5. P. 7–14.
18. Kim S.-W., Gil J.-M. Research paper classification systems based on TF-IDF and LDA schemes // *Human-Centric Computing and Information Sciences*. 2019. No 9. Article no 30. DOI: 10.1186/s13673-019-0192-7.
19. Hersh W.R., Monarch I.A., Lefferts R.G., Handerson S.K. Automatic indexing of abstracts via natural-language processing using a simple thesaurus / D.A. Evans [et al.] // *Medical Decision Making*. 1991. Vol. 11. No 4 (suppl.). P. 108–115.
20. Medelyan O., Witten I.H. Thesaurus based automatic keyphrase indexing // 6th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL '06). Chapel Hill, NC, USA, 11–15 June 2006. P. 296–297. DOI: 10.1145/1141753.1141819.
21. Голицына О.Л., Максимов Н.В., Федорова В.А. К определению семантической близости на основе связей объединенного тезауруса // Научно-техническая информация. Серия 2: Информационные процессы и системы. 2016. №. 6. С. 30–44.
22. Loukachevitch N., Nokel M., Ivanov K. Combining thesaurus knowledge and probabilistic topic models // 6th International Conference on Analysis of Images, Social Networks, and Texts (AIST 2017). Moscow, Russia, 27–29 July 2017. P. 59–71.
23. Andrzejewski D., Zhu X., Craven M. Incorporating domain knowledge into topic modeling via Dirichlet forest priors // 26th Annual International Conference on Machine Learning (ICML 2009). Montreal, Canada, 14–18 June 2009. P. 25–32. DOI: 10.1145/1553374.1553378.
24. Loukachevitch N., Ivanov K. Evaluating thesaurus-based topic models // 23rd International Conference on Applications of Natural Language to Information Systems (NLDB 2018). Paris, France, 13–15 June 2018. P. 364–376. DOI: 10.1007/978-3-319-91947-8_38.
25. ГОСТ Р ИСО 704-2010. Терминологическая работа. Принципы и методы. М.: Стандартинформ, 2012.
26. Бунин М.С., Пирумова Л.Н. Информационно-поисковый тезаурус по сельскому хозяйству и продовольствию Центральной научной сельскохозяйственной библиотеки // Российская сельскохозяйственная наука. 2020. № 5. С. 72–75. DOI: 10.31857/S2500262720050178.
27. Aubin S., Hamon T. Improving term extraction with terminological resources // 5th International Conference on Natural Language Processing (FinTAL 2006). Turku, Finland. 23–25 August 2006. P. 380–387.

Об авторах**Краснов Федор Владимирович**

кандидат технических наук;

эксперт департамента информационных систем управления, NAUMEN R&D, 620028, г. Екатеринбург, ул. Татищева, д. 49А;

E-mail: fkrasnov@naumen.ru

ORCID: 0000-0002-9881-7371

Смазневич Ирина Сергеевна

бизнес-аналитик департамента информационных систем управления, NAUMEN R&D, 620028, г. Екатеринбург, ул. Татищева, д. 49А;

E-mail: ismaznevich@naumen.ru

ORCID: 0000-0002-5996-4635

Баскакова Елена Николаевна

ведущий системный аналитик департамента информационных систем управления, NAUMEN R&D, 620028, г. Екатеринбург, ул. Татищева, д. 49А;

E-mail: enbaskakova@naumen.ru

ORCID: 0000-0002-7071-8961

The problem of loss of solutions in the task of searching similar documents: Applying terminology in the construction of a corpus vector model

Fedor V. Krasnov

E-mail: fkrasnov@naumen.ru

Irina S. Smaznevich

E-mail: ismaznevich@naumen.ru

Elena N. Baskakova

E-mail: enbaskakova@naumen.ru

NAUMEN R&D

Address: 49A, Tatishcheva Street, Ekaterinburg 620028, Russia

Abstract

This article considers the problem of finding text documents similar in meaning in the corpus. We investigate a problem arising when developing applied intelligent information systems that is non-detection of a part of solutions by the TF-IDF algorithm: one can lose some document pairs that are similar according to human assessment, but receive a low similarity assessment from the program. A modification of the algorithm, with the replacement of the complete vocabulary with a vocabulary of specific terms is proposed. The addition of thesauri when building a corpus vector model based on a ranking function has not been previously investigated; the use of thesauri has so far been studied only to improve topic models. The purpose of this work is to improve the quality of the solution by minimizing the loss of its significant part and not adding “false similar” pairs of documents. The improvement is provided by the use of a vocabulary of specific terms extracted from the text of the analyzed documents when calculating the TF-IDF values for corpus vector representation. The experiment was carried out on two corpora of structured normative and technical documents united by a subject: state standards related to information technology and to the field of railways. The glossary of specific terms was compiled by automatic analysis of the text of the documents under consideration,

and rule-based NER methods were used. It was demonstrated that the calculation of TF-IDF based on the terminology vocabulary gives more relevant results for the problem under study, which confirmed the hypothesis put forward. The proposed method is less dependent on the shortcomings of the text layer (such as recognition errors) than the calculation of the documents' proximity using the complete vocabulary of the corpus. We determined the factors that can affect the quality of the decision: the way of compiling a terminology vocabulary, the choice of the range of n -grams for the vocabulary, the correctness of the wording of specific terms and the validity of their inclusion in the glossary of the document. The findings can be used to solve applied problems related to the search for documents that are close in meaning, such as semantic search, taking into account the subject area, corporate search in multi-user mode, detection of hidden plagiarism, identification of contradictions in a collection of documents, determination of novelty in documents when building a knowledge base.

Key words: similarity of documents; semantic proximity; thesauri application; corpus vector model; applied intelligent information systems; algorithm explainability; similarity evaluation; text mining.

Citation: Krasnov F.V., Smaznevich I.S., Baskakova E.N. (2021) The problem of loss of solutions in the task of searching similar documents: Applying terminology in the construction of a corpus vector model. *Business Informatics*, vol. 15, no 2, pp. 60–74. DOI: 10.17323/2587-814X.2021.2.60.74

References

1. Krasnov F.V., Smaznevich I.S. (2020) The explicability factor of the algorithm in the problems of searching for the similarity of text documents. *Computational Technologies*, vol. 25, no 5, pp. 107–123 (in Russian). DOI: 10.25743/ICT.2020.25.5.009.
2. Otradnov K.K., Zhukov D.O., Novikova O.A. (2017) Clustering model of low-structured text data. *Modern Information Technologies and IT-Education*, vol. 13, no 3. Available at: <http://sitito.cs.msu.ru/index.php/SITITO/article/view/295> (in Russian). DOI: 10.25559/SITITO.2017.3.439.
3. Calzolari N. (1977) An empirical approach to circularity in dictionary definitions. *Cahiers de Lexicologie*, vol. 31, no 2, pp. 118–128.
4. Amsler R.A. (1980) *The structure of the Merriam-Webster pocket dictionary*. Austin, TX: The University of Texas.
5. Grefenstette G. (1994) *Explorations in automatic thesaurus discovery*. New York: Springer Science & Business Media. DOI: 10.1007/978-1-4615-2710-7.
6. Bullinaria J.A. (2008) Semantic categorization using simple word co-occurrence statistics. Proceedings of the *ESSLLI Workshop on Distributional Lexical Semantics (ESSLLI 2008)*, Hamburg, Germany, 4–9 August 2008, pp. 1–8.
7. Nazar R., Vivaldi J., Wanner L. (2012) Co-occurrence graphs applied to taxonomy extraction in scientific and technical corpora. *Procesamiento del Lenguaje Natural*, no 49, pp. 67–74.
8. Santus E., Lenci A., Lu Q., Walde S.S. (2014) Chasing hypernyms in vector spaces with entropy. Proceedings of the *14th Conference of the European Chapter of the Association for Computational Linguistics, Gothenburg, Sweden, 26–30 April 2014*, pp. 38–42.
9. Jones K.S. (1972) A statistical interpretation of term specificity and its application in retrieval. *Journal of Documentation*, vol. 28, no 1, pp. 11–21.
10. Jones K.S. (1973) Index term weighting. *Information Storage and Retrieval*, vol. 9, no 11, pp. 619–633.
11. Salton G., Yang C.S. (1973) *On the specification of term values in automatic indexing*. Ithaca, NY: Cornell University.
12. Salton G., Wong A., Yang C.S. (1975) A vector space model for automatic indexing. *Communications of the ACM*, vol. 18, no 11, pp. 613–620.
13. Otradnov K.K., Raev V.K. (2018) Experimental study of text documents vectorization techniques and their clustering algorithms efficiency. *Vestnik of Ryazan State Radio Engineering University*, no 64, pp. 73–84 (in Russian). DOI: 10.21667/1995-4565-2018-64-2-73-84.
14. Bafna P., Pramod D., Vaidya A. (2016) Document clustering: TF-IDF approach. Proceedings of the *2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT 2016)*, Chennai, India, 3–5 March 2016, pp. 61–66. DOI: 10.1109/ICEEOT.2016.7754750.
15. Fomin S.A., Belousov R.L. (2017) Detecting semantic duplicates in short news items. *Business Informatics*, no 2, pp. 47–56. DOI: 10.17323/1998-0663.2017.2.47.56.
16. Qaiser S., Ali R. (2018) Text mining: Use of TF-IDF to examine the relevance of words to documents. *International Journal of Computer Applications*, vol. 181, no 1, pp. 25–29. DOI: 10.5120/ijca2018917395.
17. Krasnov F., Ushmaev O. (2018) Exploration of hidden research directions in oil and gas industry via full text analysis of OnePetro digital library. *International Journal of Open Information Technologies*, vol. 6, no 5, pp. 7–14.
18. Kim S.-W., Gil J.-M. (2019) Research paper classification systems based on TF-IDF and LDA schemes. *Human-Centric Computing and Information Sciences*, no 9, article no 30. DOI: 10.1186/s13673-019-0192-7.
19. Evans D.A., Hersh W.R., Monarch I.A., Lefferts R.G., Handerson S.K. (1991) Automatic indexing of abstracts via natural-language processing using a simple thesaurus. *Medical Decision Making*, vol. 11, no 4 (suppl.), pp. 108–115.
20. Medelyan O., Witten I.H. (2006) Thesaurus based automatic keyphrase indexing. Proceedings of the *6th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL '06)*, Chapel Hill, NC, USA, 11–15 June 2006, pp. 296–297. DOI: 10.1145/1141753.1141819.
21. Golitsyna O.L., Maksimov N.V., Fyodorova V.A. (2016) On the definition of semantic proximity based on the links of the combined thesaurus. *Scientific and Technical Information. Series 2: Information Processes and Systems*, no 6, pp. 30–44 (in Russian).

22. Loukachevitch N., Nokel M., Ivanov K. (2017) Combining thesaurus knowledge and probabilistic topic models. Proceedings of the *6th International Conference on Analysis of Images, Social Networks, and Texts (AIST 2017)*, Moscow, Russia, 27–29 July 2017, pp. 59–71.
23. Andrzejewski D., Zhu X., Craven M. (2009) Incorporating domain knowledge into topic modeling via Dirichlet forest priors. Proceedings of the *26th Annual International Conference on Machine Learning (ICML 2009)*, Montreal, Canada, 14–18 June 2009, pp. 25–32. DOI: 10.1145/1553374.1553378.
24. Loukachevitch N., Ivanov K. (2018) Evaluating thesaurus-based topic models. Proceedings of the *23rd International Conference on Applications of Natural Language to Information Systems (NLDB 2018)*, Paris, France, 13–15 June 2018. P. 364–376. DOI: 10.1007/978-3-319-91947-8_38.
25. Standardinform (2012) *GOST R ISO 704-2010. Terminology work. Principles and methods*. Moscow: Standardinform (in Russian).
26. Bunin M.S., Pirumova L.N. (2020) Information and search thesaurus on agriculture and food of the Central Scientific Agricultural Library, *Russian Agricultural Science*, no 5, pp. 72–75 (in Russian). DOI: 10.31857/S2500262720050178.
27. Aubin S., Hamon T. (2006) Improving term extraction with terminological resources. Proceedings of the *5th International Conference on Natural Language Processing (FinTAL 2006)*, Turku, Finland, 23–25 August 2006, p. 380–387.

About the authors

Fedor V. Krasnov

Cand. Sci. (Tech.);

Expert, Department of Management Information Systems, NAUMEN R&D, 620028, 49A, Tatishcheva Street, Ekaterinburg 620028, Russia;

E-mail: fkrasnov@naumen.ru

ORCID: 0000-0002-9881-7371

Irina S. Smaznevich

Business Analyst, Department of Management Information Systems, NAUMEN R&D, 620028, 49A, Tatishcheva Street, Ekaterinburg 620028, Russia;

E-mail: ismaznevich@naumen.ru

ORCID: 0000-0002-5996-4635

Elena N. Baskakova

Leading System Analyst, Department of Management Information Systems, NAUMEN R&D, 620028, 49A, Tatishcheva Street, Ekaterinburg 620028, Russia;

E-mail: enbaskakova@naumen.ru

ORCID: 0000-0002-7071-8961

How do information technology and knowledge management affect SMEs' responsiveness to the coronavirus crisis?

Naief G. Azyabi 

E-mail: nazyabi@jazanu.edu.sa

Jazan University
Address: Jazan 45142, Saudi Arabia

Abstract

All organizations have been affected by the coronavirus pandemic in different ways. Small and medium-sized enterprises (SMEs) are more vulnerable to changes due to their limited resources. However, the capabilities of information technologies and processes of knowledge management can assist these enterprises to survive and respond appropriately to changes. Thus, this study aims to assess the extent to which information technology capabilities influence the responsiveness of SMEs to challenges that have emerged during the coronavirus crisis. It also investigates the degree to which knowledge management affects such a relationship in the context of Saudi Arabia. The study includes developing a survey as a data collection method. The responses from 136 SMEs were used to make an analysis and, consequently, draw a conclusion. It has been found that IT capabilities positively influence SMEs' responsiveness to changes brought by coronavirus, through supporting work flexibility and providing a wide range of options in the supply chain, processes, and sales. It further found that knowledge management mediates the relationship between IT capabilities and SMEs' responsiveness.

Key words: IT capabilities; small and medium-sized enterprises (SMEs); responsiveness; knowledge management; Saudi Arabia.

Citation: Azyabi N.G. (2021) How do information technology and knowledge management affect SMEs' responsiveness to the coronavirus crisis? *Business Informatics*, vol. 15, no 2, pp. 75–90. DOI: 10.17323/2587-814X.2021.2.75.90

Introduction

There are numerous ways that the coronavirus pandemic (COVID-19) affects the economy, particularly for Small and Medium-Sized Enterprises (SMEs). SMEs experience challenges such as limited resources can hinder the investment in top-notch information technology (IT) systems for efficient operations [1]. In terms of revenue, the abrupt and drastic loss of competition and profits for SMEs severely impacts their operating ability or generates a significant cash deficit. Also, the loss of jobs, concern about contagion, and increased insecurity are causing consumers to lower spending and consumption. This is compounded by the laying-off of employees and businesses unable to pay salaries. More commonly than other companies, SMEs are likely to suffer from social distancing [2].

SMEs contribute significantly to economic progress in any given country, particularly, the developing nations. In Saudi Arabia, the bulk of the companies are SMEs, with added value and employment. However, the prevalence of SMEs is more significant in some regions and sectors that have been impacted. The number of suppliers to SMEs also is lower. Furthermore, the vulnerability of SMEs increases by relying on suppliers from countries and regions with more COVID-19 [3]. Based on a decrease in global demand for their goods and services, businesses, including SMEs, will bear the brunt.

Moreover, it could be more difficult for SMEs to obtain information not only about steps to avoid the spread of the virus but also about potential business strategies to mitigate the shock and policy support initiatives. These are some severe challenges that SMEs are encountering in the midst of the pandemic, as well as in the post-coronavirus period. Since SMEs have different conditions, countries have adopted initiatives to support them. In particular, in this challenging period, several countries are urgently adopting policies to help SMEs and self-employed workers, with a heavy emphasis on short-term liquidity initiatives.

Research on the responses of SMEs to such a major and comprehensive pandemic remains limited. Thus, the need is still high to understand the behaviour of this type of organization, and how they use IT to survive during this crisis and minimize its effects. This research is considered a contribution on both theoretical and practical sides. In terms of theoretical contribution, it provides a framework that combines the effects of both IT and KM in enhancing SMEs' capabilities and responsiveness. Such a framework

will be useful and adaptable in studying or explaining other capabilities of SMEs. The contribution of this research is represented by exploring the ability of SMEs in the Kingdom of Saudi Arabia to adapt and cope with the pressures of the coronavirus pandemic. Furthermore, the contribution is presented through developing and testing a framework that can:

- ◆ provide a better understanding of IT utilization among SMEs;
- ◆ explain the effect of IT and KM processes on SMEs responses to crises.

On the practical side, this framework can guide SMEs to the best way to handle crises through proper utilization of IT and good management of knowledge.

Understanding how IT can be utilized by SMEs to survive and respond to changes could contribute to both SMEs and IT literature. Moreover, Research on SMEs' utilization of IT and knowledge management is limited in the Saudi context. Thus, this research is an attempt to fill such a gap.

This research aims to investigate the relationship between IT capabilities and Saudi SMEs' responsiveness to the coronavirus crisis. It further explores the impacts of knowledge management (KM) processes as a mediating factor of such a relationship. This research is an attempt to answer the following questions:

- ◆ To what extent do IT capabilities influence SMEs' responsiveness to the coronavirus crisis?
- ◆ To what extent do KM processes mediate such influence?

1. Literature review

1.1. IT and responsiveness capabilities in SMEs

Research has made a significant contribution to understanding the impact of the environment on IT systems [4]. However, there are gaps in the approaches that organizations, specifically SMEs, use to respond to changes in the external environment. The difference consequently raises interest in the need to understand whether IT systems have any role in supporting the SMEs in addressing viable responses in critical moments, such as the presence of a coronavirus pandemic. The relationship between IT capabilities and the level of readiness to respond to various changes in the environment has leaned more on the large enterprises [5]. The understanding of SMEs' response to environmental changes is pertinent for the fact that SMEs are more vulnerable to environmental changes

partly due to the limited resources [1]. In order to address such a gap, there is a need to assess how SMEs utilize their IT capabilities to minimize the impacts of environmental changes on them.

According to [6], “dynamic groups of ICT tools, competencies, and expertise, which are exercised through business processes, enable firms to organize activities and leverage ICT assets to achieve the desired results.” IT-based capacity is “unique.” Companies build capacity to react and organizational characteristics to environmental stimuli.

There are “internally-oriented” (IO) and “externally-oriented” (EO) capabilities [6]. IO technologies are focused on using IT applications for storing information about the internal activities of organizations, such as Enterprise Resource Planning (ERP) solutions. This method can help managers make short-term organizational choices and thereby increase performance and operational control by incorporating data through functions [7]. While technology such as ERP matures, statistics show that practices and IO-based resources among SMEs remain limited.

EO’s ability relies on IT to encourage business model creativity, new product development operations, CRM processes, and e-commerce initiatives and to enable businesses to adapt to market changes. The ability to use IT to understand and process external business knowledge is also evident in EO’s capacity, which allows companies to reconfigure their products, business models, and supply chain relationships more effectively [6]. More precisely, CRM systems support substantial market intelligence activities. IT supports the new product development, which is critical for customer participation in product innovation, and for supply chain partnership to shorten time-to-market. Because EO resources can also include the opportunity to manage e-commerce programs, the costs of reaching new business segments are controlled. Given the value of IT capabilities, even without the “weaving” of IT, SMEs will achieve strategic agility.

However, the role of these innovations in organizational processes and lowering entry costs in the new market segments can be crucial to their growth by supporting the capacities of the small companies. It can be challenging to implement ERP systems in SMEs, due to their lack of management experience in re-engineering the business process [8], and to their specific characteristics in production management, resulting in profound changes in their organizational routine and modules of standardized ERP software. Given the greater alignment

in the concept of IT requirements that they must adopt with partners, IT programs helping EO companies are challenging to implement for SMEs. As such, it requires high levels of social sophistication to apply IT in EO organizational processes. Furthermore, travel dependencies are more evident as the use of IT in product creation, and CRM requires a pre-existing build-up of IS supporting internal operations and customer profile and transaction information repositories [6].

1.2. IT and KM processes in SMEs

Through a reduction in lead times, reduction of overhead cost and product differentiation, KM aims to gain a competitive edge. Firstly, by assessing current circumstances and combining them with previous experience, KM allows companies to reduce their lead times rapidly by recognizing and responding to the evolving marketplace [9]. Secondly, by enhancing commodity quality, which eventually helps consumers, the organization aims to reduce its costs [10]. Third, KM is considered a critical enabler to improve decision making and product changes for organizational advancement [11]. To develop and sustain diverse capabilities, the productivity and effectiveness of information uprooted within an enterprise are essential [12].

KM systems are, therefore, built and implemented to achieve the above three objectives. Sher and Lee [13] have identified three issues in their previous work on KM systems, comprehensive IT design, the degree to which corporate information is built up and preserved, and search for awareness, development and dissemination of IT.

In acquiring and obtaining core expertise, IT is a vital element for organizations, whether adopting personalization or codification KM strategy. KM continues to emphasize both directly and tacitly the incorporation of core business expertise and its importance within the company. In addition to controlling and paying for private information as a central consideration for the KM system, KM takes into account the external aspect [14]. In order to ensure the convergence of the process and information gain, the internal administration role must be combined with external information through the KM system. In a KM report, the following concepts were identified: facilitation of the information, production and information acquisition, marketing knowledge production, marketing knowledge acquisition and alignment with company strategy [15].

The value of managing external knowledge is evident from literature as a specific challenge to achieve the sustainability of competitive advantage more efficiently and effectively than others [16]. It can be argued that the control of external knowledge through IT is indispensable. Furthermore, the management of external knowledge plays a significant role in achieving corporate performance and influencing the company's strategic path. To achieve such a goal, there is a need for an effective compilation of all needed information to design and execute organizational activities in alignment with corporate strategy.

1.3. KM and responsiveness capabilities of SMEs

Roach, Ryman and Makani [17] argue that for most national economies, SMEs and business enterprises are the leading segment drivers. The strategic advantage of capable SMEs is identical, allowing them to create a market niche by changing their product mix to suit the clients' needs [18]. In different parts of the world, SMEs are defined in different ways. Some define them as properties, while others use the criterion of employment, shareholder funds or sales. Some others use the hybrid criterion of a mix of income and jobs. Current literature indicates that SMEs can be distinguished by a variety of essential characteristics from larger firms. This view has now been illustrated by some international business and economic analysis, which proposes that more substantial organizations, for example, from foreign parents, may obtain less knowledge internally than from smaller organizations, because of their ability to produce information on their own [19].

Moreover, the findings discovered in Eastern Europe by Sinani and Meyer's [20] empirical research reveals that relatively smaller organizations are more knowledgeable than larger entities. Rahayu and Day [21] state that SMEs are more likely than large companies to look for information through experiences with the know-how possessors. Their specific characteristics – management structures, markets, processes, community – are significant features that influence KM activities in SMEs that separate them from large organizations.

Various variables, such as the location of the customer order point, set-up times, manufacturing resources, and system configuration (equipment and workers), affect the responsiveness of a production system. In this regard, three aspects of responsiveness are identified: product, volume and process [22].

2. Conceptual model and research hypotheses

This research involves developing a framework that is used to explore how IT capabilities and KM affected SMEs' response to changes caused by the coronavirus pandemic. The hypotheses (*Figure 1*) indicate that SMEs' responsiveness to change is affected directly by IT (independent variable), and KM is a mediating factor for this effect.

The hypotheses of this research are listed below:

- ◆ **H1:** IT capabilities positively affect SMEs' responsiveness to change;
- ◆ **H2:** IT capabilities positively affect KM processes;
- ◆ **H3:** KM processes positively affect SMEs' responsiveness to change.

2.1. IT capabilities

Zhang and Ziegelmayer [23] assert that every organization that seeks to survive change must be responsive. Flexibility in the IT infrastructure provides the relevant framework to respond to various changes in contemporary society. A firm's responsiveness refers to its ability to react quickly to the possible variations within its environment to try and seize the available opportunities [23]. The effectiveness and efficiency is a reflection of an organization's capacity to sense, internalize and act on stimuli. This is one of the processes that ensure a competitive advantage. The firms with a high level of responsiveness can outperform their competitors by utilizing the available resources to satisfy the needs of customers.

IT capabilities refer to the ability of a firm to install, utilize, integrate and reconfigure the resources from IT systems to enhance the strategies of a business [24]. This kind of approach assists businesses to catalyse their processes and support their strategy. Various approaches can lead organizations to diverse outcomes. Three types of IT capabilities have been identified: outside-in, inside-out and spanning [25]. The three subcomponents (items) can be used to measure IT capabilities. Outside-in capabilities are the external orientation that represents the ability of the firm to develop IT-links with significant business partners. Inside-out IT capabilities refer to the ability of a firm to internally deploy data, networks, as well as relevant infrastructure for its services and applications [26]. Spanning IT capabilities are the ability to integrate both the external and internal IT related factors to generate the best performance for a firm [25]. The research adopted a parsimonious yet comprehensive view of the IT capabilities that include: outside-in, inside-out and spanning.

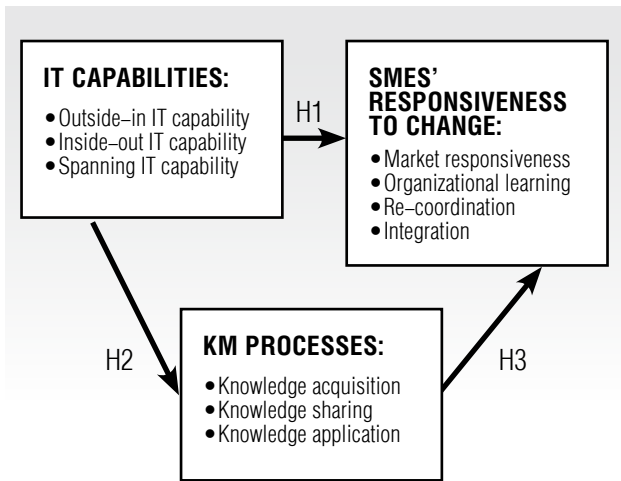


Fig. 1. Research model

The level of preparing IT personnel for unexpected changes is linked directly to the firm's plans and experience in managing such changes. The competency of such skilled personnel creates a positive impact on the responsiveness of an organization [27].

The present study would consequently seek to bridge the gap by assessing the validity of the statement. This leads to formulating the following hypothesis:

H1: IT capabilities are positively related to SMEs' responsiveness to change.

2.2. KM processes

Companies with more significant KM potential are learning to improve their ability to minimize duplication, react rapidly to change, and generate new ideas and creativity [28]. KM processes are measured by three sub-components, which are knowledge acquisition, knowledge sharing and knowledge application.

Knowledge acquisition can be conceptualized as the process involved in gaining knowledge, or only "the method of gaining information" [29]. Information acquired can be implicit, explicit or both. External networks of an organization may be an effective means of acquiring knowledge. Also, formalized and undefined external sources of knowledge provide useful information and promote developments in knowledge [30]. Knowledge sharing is a culture of social interaction involving the exchange of information, experiences and skills between the entire department or organization. The exchange of knowledge involves several collective insights that enable employees to have access to information and to develop and use networks of knowledge inside organizations [31]. Knowledge

application is the ability of employees to use the information to develop structures to resolve issues and address problems within an organization and can simply be described as information-application capabilities. Through effective use of information, individuals may make fewer errors or improve their effectiveness and minimize redundancies [12].

IT capabilities enhance the learning process by improving information acquisition, assimilation and implementation by processing contextual intelligence [32]. That is why this skill makes it possible to make the learning cycle in KM a company as an intellectual resource, which is the foundation for organizational responsiveness. IT ability may strengthen the response, as it encourages the assimilation of information in the process of learning. The IT capabilities of this sort reflect the company's ability to implement cross-functional IT frameworks, such as systems of KM and cross-company interpreting, to integrate external understanding into internal processes. This skill involves both internally and externally oriented emphasis on promoting the assimilation of information in an organization [33]. In order to test these concepts, the following hypothesis has been framed:

H2: IT capabilities positively affect KM processes.

2.3. SMEs' responsiveness to change

SMEs form an integral factor in modern economies. They form a large proportion of all the enterprises present in any given nation. However, they suffer from limitations of resources, which is imperative to respond to the various changes [34]. One of the precarious survival capabilities is responsiveness strategies. In light of the reported research study carried out by Sui and Baum [27] on the failure of more than half of enterprises, it is conceivable that SMEs' ability to adapt to changes is critical for their survival. Firms must always seek to learn from their partners that have more exceptional experiences in terms of responding to crises and market changes [35]. Knowledge plays a crucial role in impacting efforts aimed at supporting an organization to respond to various changes within its environment [36]. KM stems from the effective use of strategies and mechanisms that can ensure the processing of information and data coupled with the utilization of IT and the capacity, in terms of innovation and creativity, from the human beings [37].

Firms usually establish strength to be competitive through finding access to external knowledge. Effective firms' learning processes can support firms' responsive-

ness. The acquisition of knowledge is likely to enable a firm to respond to prevailing changes, market demands, or changes. A close coordinative relationship with the various partners can also assist an organization to get in-depth, private and rich content that they are likely to use to learn how to respond to the multiple changes in the environment. Integration with other partners can assist an organization in learning collectively on how to handle pertinent changes [38]. Coordinating with different firms can help in creating collective cognitive values. This research investigates the SMEs' responsiveness based on four dimensions: Market Responsiveness, organizational learning, re-coordination and integration. Consequently, the following hypothesis has been articulated:

H3: KM processes positively affect SMEs' responsiveness to change.

3. Research methodology

The research adopted a quantitative method of study in which a survey has been used as a data collection instrument. As stated in the introduction section, research questions were formulated to investigate the relationship between IT capabilities, KM processes and Saudi SMEs' responsiveness to the coronavirus crisis. To develop the research framework and instrument, the related literature was reviewed and analysed. Then, the elements for each framework's construct were adopted from previous research. *Table 1* shows the main literature from which the survey items were adopted.

Table 1.

Sources of survey's items

Constructs	References
Knowledge management	Yee-Loong et al., 2014 [39]
Information technologies capabilities	Zhang and Ziegelmayer, 2008 [23]
SMEs' responsiveness	Nidumolu and Knots, 1998 [40] Pavlou and El Sawy, 2006 [41]

The questionnaire consists of 30 items distributed on three main constructs as shown in the *Table 2*.

The research was conducted in the Saudi context, and the parameter which was used in SMEs' definition in this research was the number of employees (organizations with less than 200 employees). Items were translated into clear Arabic to enable targeted participants to comprehend the meaning completely. The targeted participants were asked to show their level of agreement on a five-point Likert scale regarding three main con-

structs: IT capabilities, SMEs' responsiveness to change and KM processes. The five-point Likert was 1 "strongly disagree," and 5 "strongly agree." SME executives/owners are the main targeted participants. Due to coronavirus lockdown, an online survey method was used to contact the targeted SMEs and collect the required data. All contact details of SMEs were obtained from the Chamber of Commerce without any identifying data. There were 136 SMEs who responded out of 500 targeted SMEs, which counts for 27.2% as response rate.

The questionnaire elements were tested for their reliability and robustness before the key findings are expanded. Cronbach's Alpha value is calculated to determine the value of the items' reliability. The factor analysis was carried out to find out the validity of each construct and to report the value of each factor loading. The critical analyses were designed to test the hypotheses and to construct the relationship between variables, mainly via Pearson product-moment correlation and multiple regressions. In the beginning, a general description of the data collected was given and descriptive statistics such as frequency distribution, central pattern and variability measurements were extracted from the use of the SPSS program.

From the data collected, 136 participants responded to the survey. Out of these, the respondents were 36% Chief Executive officers, 11.8% directors of accounting and finance, 14.7% directors of human resources, 17.6% were directors of IT while others were 19.9%. The data shows that majority of the participants are in managerial positions, which puts them in a better position to answer the research questions.

The demographic analysis (*Table 3*) shows that there is a wide diversity of participant SMEs in terms of sectors (18.4% retail firms, 14% in manufacturing firms, 21.3% restaurants and food industry, 10.3% IT services, 14.7% constructions, and 21.3% in others sectors). The majority of the firms have been in operation between two to five years, have between 50–249 employees, and with annual revenue of less than 3 million Saudi Arabian Riyals.

4. Results and analysis

This research intends to investigate the relationship between IT capabilities and the SMEs' responsiveness during the coronavirus pandemic. The study also aims to investigate whether KM mediates this relationship or not. To achieve these particular investigations, correlation and regression analysis were used. There were 136 SMEs' managers who participated in the survey. The collected data was inserted into SPSS statistics software, and the research hypotheses were examined.

Table 2.

The survey's items

Main construct	Sub-constructs	Items
Knowledge management processes	Knowledge acquisition	KA1: Our organization has processes for generating new knowledge based on existing knowledge. KA2: Our organization has processes for acquiring customer knowledge. KA3: Our organization has processes for acquiring knowledge on developing new products/services.
	Knowledge sharing	KS1: Our organization has processes for sharing knowledge throughout the organization. KS2: Our organization has processes for transferring organizational knowledge to employees. KS3: Our organization has processes for sharing knowledge among business partners.
	Knowledge application	KAP1: Our organization has processes for integrating different sources and types of knowledge. KAP2: Our organization has processes for applying experiential knowledge. KAP3: Our organization has processes for applying knowledge to solve new problems.
IT capabilities	Outside-in IT capabilities	OIT1: Our organization has technology-based links with customers. OIT2: Our organization has technology-based links with suppliers. OIT3: Our organization has IT-based entrepreneurial collaborations with external partners.
	Inside-out IT capabilities	IIT1: Our organization has good appropriateness of the data architectures. IIT2: Our organization has good appropriateness of network architectures. IIT3: Our organization has good adequacy of architecture flexibility.
	Spanning IT capabilities	SIT1: Our organization restructure IT work processes to leverage opportunities. SIT2: Our organization has multidisciplinary teams to blend business and technology expertise. SIT3: Our organization has a climate that nurtures IT project championship.
SMEs' responsiveness	Market responsiveness	MR1: During the coronavirus crisis, our organization has the agility to respond to environmental changes. MR2: During the coronavirus crisis, our organization monitors the environment for the identification of new business opportunities. MR3: During the coronavirus crisis, our organization has the agility to develop new products and services.
	Organizational learning	OL1: Coronavirus crisis increases our organization's desire to improve the ability of employees. OL2: Coronavirus crisis encourages our organization to have a learning atmosphere. OL3: Coronavirus crisis encourages our organization to enhance the sharing of ideas.
	Re-coordination	RE1: During the coronavirus crisis, our organization has coordination between organizational processes and employees' abilities. RE2: During the coronavirus crisis, our organization allocates the assignment of activities to employees based on their abilities. RE3: During the coronavirus crisis, our organization engages in effective assignment of resources to projects and processes.
	Integration	INT1: During the coronavirus crisis, our organization engages in defining new activities based on new situations. INT2: During the coronavirus crisis, our organization enhances the reintegration of organizational processes. INT3: During the coronavirus crisis, there is sufficient cooperation of different departments or sections in our organization.

4.1. KMO and Barlett's Test of Sphericity

KMO's test of sampling adequacy and Barlett's test of sphericity (Table 4) were carried out. KMO's value is 0.955, with a p -value of 0.000 (< 0.05). The Barlett's test of sphericity is also significant ($p = 0.000$; < 0.05).

Principal component analysis can consequently be conducted. The factor loading for the various elements can be as shown in Table 5.

Rahn [42] states that a factor loading of more than 0.4 is good, and a researcher can carry out further analysis. From the results in Table 5, all items have factor loadings of more than 0.4 and consequently qualify for further analysis.

Table 3.

Demographic analysis

SMEs' sector		Role of respondents	
Sector	%	Role	%
Retail	25 (18.4%)	Chief Executive Officer	49 (36%)
Manufacturing	19 (14%)	Director of Finance	16 (11.8%)
Restaurants and food industry	29 (21.3%)	Director of HR	20 (14.7%)
IT services	14 (10.3%)	Director of IT	24 (17.6%)
Constructions (real estate, engineers)	20 (14.7%)	Other	27 (19.9%)
Others	29 (21.3%)		

SMEs' age		Number of employees		Annual revenue	
Age	%	No of employees	%	Revenue	%
< 2 years	34 (25%)	1–5	42 (30.9%)	< 3 million	58 (42.6%)
2–5 years	41 (30.1%)	6–49	44 (32.4%)	3–40 million	40 (29.4%)
5–10 years	24 (17.6%)	50–249	50 (36.8%)	40–200 million	38 (27.9%)
> 10 years	37 (27.2%)				

4.2. Cronbach's Alpha

Cronbach's Alpha was used to measure the level of reliability of the items in the scale for every variable. It measures the level of interrelatedness or the homogeneity of the items [43]. It can be used to assess how a group of given items can measure a construct. The parameter would determine if the individual items would be relied on to measure the intended variables. According to Sharma [44], a reliability coefficient of ≥ 0.7 is acceptable. Table 6 shows Cronbach's Alpha for the various items and their respective constructs.

The reliability of the various items was calculated using Cronbach's Alpha, as shown in Table 6. The values are higher than 0.7, which shows higher reliability. The internal consistency of the various research question items when one item is deleted is shown in Table 7. From the results, deleting either of the items would reduce Cronbach's Alpha. Consequently, all items contribute to a higher internal consistency and reliability and can all be used for further analysis.

Table 4.

KMO's test of sampling adequacy and Barlett's test of sphericity

KMO and Bartlett's test		
Kaiser–Meyer–Olkin measure of sampling adequacy		.955
Bartlett's test of sphericity	Approx. Chi-Square	3583.843
	df	435
	Sig.	.000

Table 5.

Factor loading for the various items

Items	Factor Loading	Items	Factor Loading
KA1	0.753	SIT1	0.728
KA2	0.720	SIT2	0.780
KA3	0.785	SIT3	0.772
KS1	0.773	MR1	0.772
KS2	0.787	MR2	0.740
KS3	0.784	MR3	0.779
KAP1	0.750	OL1	0.699
KAP2	0.804	OL2	0.755
KAP3	0.810	OL3	0.748
OIT1	0.742	RE1	0.756
OIT2	0.765	RE2	0.721
OIT3	0.781	RE3	0.737
IIT1	0.757	INT1	0.763
IIT2	0.775	INT2	0.744
IIT3	0.769	INT3	0.714

4.3. Correlation analysis

Correlation analysis was carried out to test the identified hypotheses.

H1: IT capabilities positively affect SMEs' responsiveness to change (Table 8).

The analysis shows a strong positive correlation between KM processes and SMEs' responsiveness ($r = 0.838$; $p < 0.05$). The null hypothesis is, therefore, rejected.

H2: IT capabilities positively affect KM processes (Table 9).

The analysis shows a strong positive correlation between IT capabilities and KM processes ($r = 0.882$; $p < 0.05$).

H3: KM processes positively affect SMEs' responsiveness to change (Table 10).

The results show that there is a strong positive correlation between IT capabilities and SMEs' responsiveness ($r = 0.832$; $p < 0.05$).

Table 6.

Cronbach's Alpha for the variables

Variable	Number of items	Cronbach's Alpha
IT capabilities	9	0.937
KM processes	9	0.931
SMEs' responsiveness	12	0.943

Table 7.

Cronbach's Alpha if items deleted

KM processes		IT capabilities		SMEs' responsiveness	
Items	Cronbach's Alpha if item deleted	Items	Cronbach's Alpha if item deleted	Items	Cronbach's Alpha if item deleted
KA1	.926	OIT1	.930	MR1	.939
KA2	.923	OIT2	.932	MR2	.938
KA3	.923	OIT3	.931	MR3	.938
KS1	.923	IIT1	.929	OL1	.940
KS2	.921	IIT2	.927	OL2	.938
KS3	.922	IIT3	.930	OL3	.938
KAP1	.926	SIT1	.933	RE1	.937
KAP2	.921	SIT2	.931	RE2	.938
KAP3	.923	SIT3	.928	RE3	.939
				INT1	.938
				INT2	.938
				INT3	.939

4.4. Regression analysis

A regression analysis was carried out to determine the extent of IT capabilities' influence on SMEs' responsiveness. The results of the analysis are shown in Table 11.

From the analysis shown in the Table 11, IT capabilities are proved to have a strong positive correlation between IT capabilities and SMEs' responsiveness (R-value 0.832; $P < 0.05$); $F(1,134) = 301.158$. The β -value is 0.832, $P = 0.000$, which is as well as a significant value. The R-squared value (0.692) implies that the IT capabilities can account for 69.2% of the variation in the SMEs' efforts to respond to the impacts instigated by the coronavirus pandemic. The other factors would only account for 31.8% of the variation. The prediction model for the analysis is the following:

$$\text{SMEs' responsiveness} = 0.772 \cdot \text{IT capabilities} + 0.931.$$

Table 8.

Correlation analysis between KM and SMEs' responsiveness

		Knowledge management	SMEs' responsiveness
Knowledge management	Pearson Correlation	1	.838**
	Sig. (2-tailed)		.000
	N	136	136
SMEs' responsiveness	Pearson Correlation	.838**	1
	Sig. (2-tailed)	.000	
	N	136	136

** Correlation is significant at the 0.01 level (2-tailed)

Table 9.

Correlation analysis between IT capabilities and knowledge management

		IT capabilities	KM processes
IT capabilities	Pearson Correlation	1	.882**
	Sig. (2-tailed)		.000
	N	136	136
KM processes	Pearson Correlation	.882**	1
	Sig. (2-tailed)	.000	
	N	136	136

** Correlation is significant at the 0.01 level (2-tailed)

Table 10.

**Correlation between IT capabilities
and SMEs' responsiveness**

		IT capabilities	SMEs' responsiveness
IT capabilities	Pearson Correlation	1	.832**
	Sig. (2-tailed)		.000
	N	136	136
SMEs' responsiveness	Pearson Correlation	.832**	1
	Sig. (2-tailed)	.000	
	N	136	136

** Correlation is significant at the 0.01 level (2-tailed)

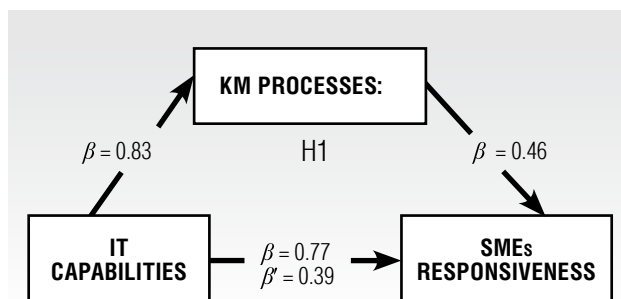


Fig. 2. Path Diagram Model for the relationship among the various variables

4.5. Mediation test using Andrew F. Hayes PROCESS method

From the results shown in Figure 2 and Tables 12 – 14, IT is a significant predictor of knowledge management, $\beta = 0.83$; ($p = 0.000$). Multiple regression shows that IT and KM processes are significant predictors of SMEs' responsiveness; $p = 0.000$, respectively. IT is a significant predictor of SMEs' responsiveness ($\beta = 0.77$; $p = 0.000$). The results also prove that the indirect effect of IT capabilities is significantly greater

than zero. For instance, in this case, the effect size is 0.38, with a 95% confidence interval, which does not include a zero. The effect is consequently greater than zero at $\alpha = 0.05$.

The first step aimed to assess if the predictor variable (IT) significantly predicts the mediator variable (KM processes). In comparison, the second step is examined if both IT and KM processes significantly predict SMEs' responsiveness. The last step is investigated if IT significantly predicts SMEs' responsiveness. The use of both the predictor and the mediator shows a significant decline in β -value, which proves a partial mediation from the intervening variable. Consequently, it can be argued that KM partially mediates the relationship between IT and the SMEs' responsiveness to the impacts of the coronavirus pandemic.

5. Discussion

Improved IT around the globe has contributed to tremendous changes in the way SMEs are operating. The presence of a coronavirus pandemic has come with its share of drawbacks with devastating consequences on the limited resources businesses such as SMEs. They need to be flexible in order to escape the challenges imposed by the ever-changing environment. Rigid businesses, in most cases, fail. One of the most significant factors to enhance organizations' flexibility is through adopting the proper IT capabilities. IT capabilities, through supporting efficient management of internal activities, can lead organizations to quick and accurate decisions. Additionally, with the dynamism brought about by the rise of the coronavirus pandemic, IT capabilities can be used to obtain information and knowledge from external sources, which can be an added value to the organization's operations. It further enhances and facilitates communication regardless of geographic restrictions, and to enable work from a distance in some tasks.

Ajayi and Olayungbo [45] affirm that IT plays a vital role in strengthening the competitive advantage of SMEs, and can be used to access the international markets. Telecommuting, for example, can be used as motivational factors for individuals to work for SMEs. The use of information technology is a motivator for young

Table 11.

Results of regressing SMEs' responsiveness against IT capabilities

Model	R	R Square	d.f	F	Coefficients	Beta	Sig. Value
1	.832	.692	(1,134)	301.158	Constant = 0.931 IT capabilities = 0.772	0.832	0.000

Table 12.

**Model summary:
IT capabilities and knowledge management**

	R	R-sq.	MSE	F	df1	df2	p
	.88	.78	.09	468.26	1.00	134.00	.00
	Coeff	se	t	p	LLCI	ULCI	
Constant	.71	.16	4.52	.00	.40	1.03	
IT_CAP	.83	.04	21.64	.00	.75	.90	

Table 13.

**Model summary:
IT capabilities, knowledge management,
and SMEs' responsiveness**

	R	R-sq.	MSE	F	df1	df2	p
	.86	.74	.10	190.05	2.00	133.00	.00
	Coeff	se	t	p	LLCI	ULCI	
Constant	.60	.18	3.31	.00	.24	.96	
IT_CAP	.39	.09	4.48	.00	.22	.56	
KNWMGT	.46	.09	5.00	.00	.28	.65	

Table 14.

Andrew F. Hayes analysis results

**Model summary:
IT capabilities and SMEs' responsiveness**

	R	R-sq.	MSE	F	df1	df2	p
	.83	.69	.12	301.16	1.00	134.00	.00
	Coeff	se	T	p	LLCI	ULCI	
Constant	.93	.18	5.07	.00	.57	1.29	
IT_CAP	.77	.04	17.35	.00	.68	.86	

The total effect of X on Y

Effect	se	T	p	LLCI	ULCI	c_ps	c_cs
.77	.04	17.35	.00	.68	.86	1.25	.83

The direct effect of X on Y

Effect	se	t	p	LLCI	ULCI	c'_ps	c'_cs
.39	.09	4.48	.00	.22	.56	.63	.42

Indirect effect(s) of X on Y:

	Effect	BootSE	BootLLCI	BootULCI
KNWMGT	.38	.09	.22	.55

employees whose firms are unable to pay higher salaries due to the impacts of the coronavirus. Using information technology allows employees to operate from home and works to curb the spread of coronavirus. SMEs can consequently minimize the number of workers in the office to allow the majority to work from home. Enterprises can, therefore, work as they follow directives from the Health authorities.

Knowledge can be acquired from clients or other firms. Sharing of knowledge with both the internal and external agencies can contribute much to supporting the SMEs to make better decisions on their various operations. This practice can be a contributing factor in dealing with the impacts of the COVID-19 pandemic.

The study also reports that KM processes act as an essential mediator between IT capabilities and the SMEs' responsiveness. The mediation is effected by three facets of KM processes that include knowledge acquisition, knowledge sharing and knowledge application. Vaidyanathan and Kidambi [31] argue that organizations need the exchange of information among employees to allow the provision of insights, develop and use networks of knowledge, particularly in SMEs. The flow of information in a firm facilitates both innovations as well as general development [29]. The use of social media information, for instance, can help firms to know the reactions of the various firms, present, and potential customers and the kind of services and products highly needed by these customers. Innovative approaches always form the archetypal way to ensure that SMEs thrive in a market characterized by unprecedented changes, challenges and opportunities.

Conclusion

The study proposed a model that linked IT and SMEs' responsiveness, IT and KM, and KM and SMEs' responsiveness to changes brought by the coronavirus pandemic. The study reveals the relationship between IT capabilities and SMEs' responsiveness. A relationship between KM processes and the SMEs' responsiveness in the wake of the coronavirus pandemic has been revealed. IT capabilities were also assessed to relate to knowledge management processes. The Andrew F. Bayes mediation test carried out ascertained that knowledge management partially mediates between IT capabilities and the SMEs' responsiveness. Information technology consequently can be used to support the operations of SMEs during changes brought about by the coronavirus pandemic. Furthermore, knowledge management processes could be used to support this relationship through the flow of pertinent information for developing the business processes. ■

References

1. Pathan Z.H., Jianqiu Z., Latif Z., Memon S., Hao D. (2017) Research on ICT patterns and integration in small and medium enterprises: A case study of Pakistan. *International Journal of u- and e-Service, Science and Technology*, vol. 10, no 2, pp. 53–64. DOI: 10.14257/ijunesst.2017.10.2.06.
2. Bénassy-Quéré A., Marimon R., Pisani-Ferry J., Reichlin L., Schoenmaker D., Weder B. (2020) *COVID-19: Europe needs a catastrophe relief plan*. Available at: <https://voxeu.org/article/covid-19-europe-needs-catastrophe-relief-plan> (accessed 11 March 2020).
3. Berger Z.D., Evans N.G., Phelan A.L., Silverman R.D. (2020) Covid-19: control measures must be equitable and inclusive. *BMJ*, 368, m1141. DOI: 10.1136/bmj.m1141.
4. Benešová D., Hu ek M. (2019) Factors for efficient use of information and communication technologies influencing sustainable position of service enterprises in Slovakia. *Entrepreneurship and Sustainability Issues*, vol. 6, no 3, pp. 1082–1094. DOI: 10.9770/jesi.2019.6.3(9).
5. Dalla Pellegrina L., Frazzoni S., Rotondi Z., Vezzulli A. (2017) Does ICT adoption improve access to credit for small enterprises? *Small Business Economics*, vol. 48, no 3, pp. 657–679. DOI: 10.1007/s11187-016-9794-x.
6. Neirotti P., Raguseo E., Paolucci E. (2018) How SMEs develop ICT-based capabilities in response to their environment. *Journal of Enterprise Information Management*, vol. 31, no 1, pp. 10–37. DOI: 10.1108/JEIM-09-2016-0158.
7. Mithas S., Rust R.T. (2016) How information technology strategy and investments influence firm performance: Conjecture and empirical evidence. *MIS Quarterly*, vol. 40, no 1, pp. 223–245. DOI: 10.25300/MISQ/2016/40.1.10.
8. Gupta H., Aye K.T., Balakrishnan R., Rajagopal S., Nguwi Y.Y. (2014) Formulating, implementing and evaluating ERP in small and medium scale industries. *International Journal of Advances in Computer Science and Technology*, vol. 3, no 6, pp. 386–389.
9. Giampaoli D., Ciambotti M., Bontis N. (2017) Knowledge management, problem solving and performance in top Italian firms. *Journal of Knowledge Management*, vol. 21, no 2, pp. 355–375. DOI: 10.1108/JKM-03-2016-0113.
10. Azyabi N.G. (2017) The role of information technology in enhancing SMEs capabilities through knowledge management. *Journal of Organizational Knowledge Management*, vol. 10, no 1, pp. 98–109.
11. Darestani H., Ismail W.K.W., Heng L.H. (2016) Investigating a new framework for the impact of knowledge management strategy on organizational performance. *Journal of Soft Computing and Decision Support Systems*, vol. 3, no 5, pp. 1–23.
12. Blome C., Schoenherr T., Eckstein D. (2014) The impact of knowledge transfer and complexity on supply chain flexibility: A knowledge-based view. *International Journal of Production Economics*, vol. 147, part B, pp. 307–316. DOI: 10.1016/j.ijpe.2013.02.028.
13. Sher P.J., Lee V.C. (2004) Information technology as a facilitator for enhancing dynamic capabilities through knowledge management. *Information & Management*, vol. 41, no 8, pp. 933–945. DOI: 10.1016/j.im.2003.06.004.
14. ul Hasan S., Macchi M., Pozzetti A., Carrasco-Gallego R. (2014) Achieving responsiveness in small and medium-sized enterprises through assemble to order strategy. Proceedings of the *IFIP International Conference on Advances in Production Management Systems* (APMS 2014), Ajaccio, France, 20–24 September 2014, pp. 208–215.
15. Mao H., Liu S., Zhang J., Deng Z. (2016) Information technology resource, knowledge management capability, and competitive advantage: The moderating role of resource commitment. *International Journal of Information Management*, vol. 36, no 6, part A, pp. 1062–1074. DOI: 10.1016/j.ijinfomgt.2016.07.001.
16. Nupap S., Chakpitak N., Neubert G., Tra-Ngarn Y. (2016) Knowledge management system for Thai small and medium-sized enterprises. *International Journal of Innovation and Learning*, vol. 19, no 2, pp. 150–168. DOI: 10.1504/IJIL.2016.074467.
17. Roach D.C., Ryman J.A., Makani J. (2016) Effectuation, innovation and performance in SMEs: an empirical study. *European Journal of Innovation Management*, vol. 19, no 2, pp. 214–238. DOI: 10.1108/EJIM-12-2014-0119.
18. Grace P., Kozak R., Nelson H. (2018) Understanding SME success in the value-added forest products sector: Insights from British Columbia. *BioProducts Business*, vol. 3, no 9, pp. 107–117. DOI: 10.22382/bbp-2018-009.
19. Mohammed A.A., Hafeez-Baig A., Gururajan R. (2019) An exploratory qualitative research to address processes that are utilised for managing knowledge: a case study in a Queensland Regional University. *International Journal of Higher Education and Sustainability*, vol. 2, no 3, pp. 173–196. DOI: 10.1504/IJHES.2019.101954.
20. Iwasaki I., Tokunaga M. (2016) Technology transfer and spillovers from FDI in transition economies: A meta-analysis. *Journal of Comparative Economics*, vol. 44, no 4, pp. 1086–1114. DOI: 10.1016/j.jce.2016.10.005.
21. Rahayu R., Day J. (2015) Determinant factors of e-commerce adoption by SMEs in developing country: evidence from Indonesia. *Procedia-Social and Behavioral Sciences*, vol. 195, pp. 142–150. DOI: 10.1016/j.sbspro.2015.06.423.
22. Belvedere V., Grando A., Papadimitriou T. (2010) The responsiveness of Italian small-to-medium sized plants: Dimensions and determinants. *International Journal of Production Research*, vol. 48, no 21, pp. 6481–6498. DOI: 10.1080/00207540903234751.
23. Zhang J., Li H., Ziegelmayer J.L. (2008) An empirical study of the relationship between IT infrastructure flexibility and IT responsiveness in SMEs: A resource-based analysis. Proceedings of the *Fourteenth Americas Conference on Information Systems*, Toronto, ON, Canada, 14–17 August 2008, pp. 1–7.
24. Cai Z., Huang Q., Liu H., Liang L. (2016) The moderating role of information technology capability in the relationship between supply chain collaboration and organizational responsiveness: Evidence from China. *International Journal of Operations and Production Management*, vol. 36, no 10, pp. 1247–1271. DOI: 10.1108/IJOPM-08-2014-0406.
25. Wade M., Hulland J. (2004) Review: the resource-based view and information systems research: review, extension, and suggestions for future research. *MIS Quarterly*, vol. 28, no 1, pp. 107–142. DOI: 10.2307/25148626.

26. Bharadwaj A.S., Sambamurthy V., Zmud R.W. (1999) IT capabilities: theoretical perspectives and empirical operationalization. *Proceedings of the 20th International Conference on Information Systems (ICIS 99), Charlotte, North Carolina, USA, 13–15 December 1999*, pp. 378–385.
27. Sui S., Baum M. (2014) Internationalization strategy, firm resources and the survival of SMEs in the export market. *Journal of International Business Studies*, vol. 45, no 7, pp. 821–841. DOI: 10.1057/jibs.2014.11.
28. Akhtar M.W., Imran M. (2015) Knowledge management process, HRM practices and individual performance. *Australian Journal of Social Science Research*, vol. 2, no 2, pp. 8–15.
29. Åkerman N. (2015) Knowledge-acquisition strategies and the effects on market knowledge – profiling the internationalizing firm. *European Management Journal*, vol. 33, no 2, pp. 79–88. DOI: 10.1016/j.emj.2014.06.003.
30. Yu X., Yan J., Assimakopoulos D. (2015) Case analysis of imitative innovation in Chinese manufacturing SMEs: Products, features, barriers and competences for transition. *International Journal of Information Management*, vol. 35, no 4, pp. 520–525. DOI: 10.1016/j.ijinfomgt.2015.03.003.
31. Vaidyanathan S., Kidambi S.S. (2018) An empirical evaluation of adoption and diffusion of new ICTs for knowledge sharing in IT organizations. *International Journal of Web Portals*, vol. 10, no 1, pp. 1–14. DOI: 10.4018/IJWP.2018010101.
32. Chakravarty A., Grewal R., Sambamurthy V. (2013) Information technology competencies, organizational agility, and firm performance: Enabling and facilitating roles. *Information systems research*, vol. 24, no 4, pp. 976–997. DOI: 10.1287/isre.2013.0500.
33. Taylor P. (2015) The importance of information and communication technologies (ICTs): An integration of the extant literature on ICT adoption in small and medium enterprises. *International Journal of Economics, Commerce and Management*, vol. 3, no 5, pp. 274–290.
34. Gancarczyk M., Gancarczyk J. (2018) Proactive international strategies of cluster SMEs. *European Management Journal*, vol. 36, no 1, pp. 59–70. DOI: 10.1016/j.emj.2017.03.002.
35. de Jesus Pacheco D.A., Carla S., Jung C.F., Ribeiro J.L.D., Navas H.V.G., Cruz-Machado V.A. (2017) Eco-innovation determinants in manufacturing SMEs: Systematic review and research directions. *Journal of Cleaner Production*, vol. 142, no 4, pp. 2277–2287. DOI: 10.1016/j.jclepro.2016.11.049.
36. Akram M.U., Chauhan C., Ghosh K., Singh A. (2019) Knowledge management, sustainable business performance and empowering leadership: A firm-level approach. *International Journal of Knowledge Management*, vol. 15, no 2, pp. 20–35. DOI: 10.4018/IJKM.2019040102.
37. Jorna R. (Ed.) (2017) *Sustainable innovation: The organisational, human and knowledge dimension*. Routledge.
38. Song J., Li F., Wu D.D., Liang L., Dolgui A. (2017) Supply chain coordination through integration of innovation effort and advertising support. *Applied Mathematical Modelling*, no 49, pp. 108–123. DOI: 10.1016/j.apm.2017.04.041.
39. Yee-Loong Chong A., Ooi K.B., Bao H., Lin B. (2014) Can e-business adoption be influenced by knowledge management? An empirical analysis of Malaysian SMEs. *Journal of Knowledge Management*, vol. 18, no 1, pp. 121–136. DOI: 10.1108/JKM-08-2013-0323.
40. Nidumolu S.R., Knotts G.W. (1998) The effects of customizability and reusability on perceived process and competitive performance of software firms. *MIS Quarterly*, vol. 22, no 2, pp. 105–137.
41. Pavlou P.A., El Sawy O.A. (2006) From IT leveraging competence to competitive advantage in turbulent environments: The case of new product development. *Information Systems Research*, vol. 17, no 3, pp. 198–227. DOI: 10.1287/isre.1060.0094.
42. Rahn R. (2014) *A comprehensive critique of the American College Health Association's National College Health Assessment survey instrument* (Doctoral dissertation).
43. Henson R.K. (2001) Teacher self-efficacy: Substantive implications and measurement dilemmas. Presented at the *Annual Meeting of the Educational Research Exchange, College Station, TX, USA, 26 January 2001*, p. 1–43.
44. Sharma B. (2016) A focus on reliability in developmental research through Cronbach's Alpha among medical, dental and paramedical professionals. *Asian Pacific Journal of Health Sciences*, vol. 3, no 4, pp. 271–278. DOI: 10.21276/apjhs.2016.3.4.43.
45. Ajayi A., Olayungbo D.O. (2014) ICT adoption in small and medium scale enterprises in Nigeria: An assessment. *International Journal of Research*, vol. 1, no 9, pp. 889–897.

About the author

Naief G. Azyabi

Ph.D.;

Associate Professor, Department of Management Information Systems, College of Business Administration, Jazan University, Jazan 45142, Saudi Arabia;

E-mail: nazyabi@jazanu.edu.sa

ORCID: 0000-0003-0798-8114

Как информационные технологии и управление знаниями влияют на восприимчивость среднего и малого бизнеса к коронавирускому кризису

Н.Г. Азяби

E-mail: nazyabi@jazanu.edu.sa

Jazan University

Address: Jazan 45142, Saudi Arabia

Аннотация

Все организации в разной степени страдают от пандемии коронавируса. При этом средние и малые предприятия уязвимы в большей степени, ввиду ограниченности их ресурсов. Однако возможности информационных технологий и процессов управления знаниями могут помочь этим предприятиям выжить и адекватно реагировать на происходящие изменения. Таким образом, целью настоящего исследования является оценка того, в какой степени возможности информационных технологий влияют на способность средних и малых предприятий реагировать на проблемы, возникающие во время вызванного коронавирусом кризиса. Также исследуется степень влияния управления знаниями на такие способности в контексте Саудовской Аравии. Исследование основано на проведении опросов в качестве метода сбора данных. Ответы 136 средних и малых предприятий использованы для анализа и формирования выводов. Было выявлено, что возможности информационных технологий положительно влияют на способность предприятий реагировать на изменения, вызванные коронавирусом, за счет поддержки гибкости работы и широкой вариативности в цепочках поставок, процессах и продажах. Кроме того, было установлено, что управление знаниями способствует взаимосвязи между возможностями информационных технологий и оперативностью реагирования предприятий на изменения.

Ключевые слова: возможности информационных технологий; средний и малый бизнес; восприимчивость; управление знаниями; Саудовская Аравия.

Цитирование: Azyabi N.G. How do information technology and knowledge management affect SMEs' responsiveness to the coronavirus crisis? // Business Informatics. Vol. 15. No 2. P. 75–90. DOI: 10.17323/2587-814X.2021.2.75.90

Литература

1. Research on ICT patterns and integration in small and medium enterprises: A case study of Pakistan / Z.H. Pathan [et al.] // International Journal of u- and e-Service, Science and Technology. 2017. Vol. 10. No 2. P. 53–64. DOI: 10.14257/ijunesst.2017.10.2.06.
2. COVID-19: Europe needs a catastrophe relief plan / A. Bénassy-Quéré [et al.], 2020. [Электронный ресурс]: <https://voxeu.org/article/covid-19-europe-needs-catastrophe-relief-plan> (дата обращения 11.03.2020).
3. Berger Z.D., Evans N.G., Phelan A.L., Silverman R.D. Covid-19: control measures must be equitable and inclusive // BMJ. 2020. 368, m1141. DOI: 10.1136/bmj.m1141.
4. Benešová D., Hušek M. Factors for efficient use of information and communication technologies influencing sustainable position of service enterprises in Slovakia // Entrepreneurship and Sustainability Issues. 2019. Vol. 6. No 3. P. 1082–1094. DOI: 10.9770/jesi.2019.6.3(9).
5. Dalla Pellegrina L., Frazzoni S., Rotondi Z., Vezzulli A. Does ICT adoption improve access to credit for small enterprises? // Small Business Economics. 2017. Vol. 48. No 3. P. 657–679. DOI: 10.1007/s11187-016-9794-x.
6. Neirotti P., Raguseo E., Paolucci E. How SMEs develop ICT-based capabilities in response to their environment // Journal of Enterprise Information Management. 2018. Vol. 31. No 1. P. 10–37. DOI: 10.1108/JEIM-09-2016-0158.
7. Mithas S., Rust R.T. How information technology strategy and investments influence firm performance: Conjecture and empirical evidence // MIS Quarterly. 2016. Vol. 40. No 1. P. 223–245. DOI: 10.25300/MISQ/2016/40.1.10.
8. Formulating, implementing and evaluating ERP in small and medium scale industries / H. Gupta [et al.] // International Journal of Advances in Computer Science and Technology. 2014. Vol. 3. No 6. P. 386–389.

9. Giampaoli D., Ciambotti M., Bontis N. Knowledge management, problem solving and performance in top Italian firms // *Journal of Knowledge Management*. 2017. Vol. 21. No 2. P. 355–375. DOI: 10.1108/JKM-03-2016-0113.
10. Azyabi N.G. The role of information technology in enhancing SMEs capabilities through knowledge management // *Journal of Organizational Knowledge Management*. 2017. Vol. 10. No 1. P. 98–109.
11. Darestani H., Ismail W.K.W., Heng L.H. Investigating a new framework for the impact of knowledge management strategy on organizational performance // *Journal of Soft Computing and Decision Support Systems*. 2016. Vol. 3. No 5. P. 1–23.
12. Blome C., Schoenherr T., Eckstein D. The impact of knowledge transfer and complexity on supply chain flexibility: A knowledge-based view // *International Journal of Production Economics*. 2014. Vol. 147. Part B. P. 307–316. DOI: 10.1016/j.ijpe.2013.02.028.
13. Sher P.J., Lee V.C. Information technology as a facilitator for enhancing dynamic capabilities through knowledge management // *Information & Management*. 2004. Vol. 41. No 8. P. 933–945. DOI: 10.1016/j.im.2003.06.004.
14. ul Hasan S., Macchi M., Pozzetti A., Carrasco-Gallego R. Achieving responsiveness in small and medium-sized enterprises through assemble to order strategy // *Proceedings of the IFIP International Conference on Advances in Production Management Systems (APMS 2014)*, Ajaccio, France, 20–24 September 2014, pp. 208–215.
15. Mao H., Liu S., Zhang J., Deng Z. Information technology resource, knowledge management capability, and competitive advantage: The moderating role of resource commitment // *International Journal of Information Management*. 2016. Vol. 36. No 6. Part A. P. 1062–1074. DOI: 10.1016/j.ijinfomgt.2016.07.001.
16. Nupap S., Chakpitak N., Neubert G., Tra-Ngarn Y. Knowledge management system for Thai small and medium-sized enterprises // *International Journal of Innovation and Learning*. 2016. Vol. 19. No 2. P. 150–168. DOI: 10.1504/IJIL.2016.074467.
17. Roach D.C., Ryman J.A., Makani J. Effectuation, innovation and performance in SMEs: an empirical study // *European Journal of Innovation Management*. 2016. Vol. 19. No 2. P. 214–238. DOI: 10.1108/EJIM-12-2014-0119.
18. Grace P., Kozak R., Nelson H. Understanding SME success in the value-added forest products sector: Insights from British Columbia // *BioProducts Business*. 2018. Vol. 3. No 9. P. 107–117. DOI: 10.22382/bpb-2018-009.
19. Mohammed A.A., Hafeez-Baig A., Gururajan R. An exploratory qualitative research to address processes that are utilised for managing knowledge: a case study in a Queensland Regional University // *International Journal of Higher Education and Sustainability*. 2019. Vol. 2. No 3. P. 173–196. DOI: 10.1504/IJHES.2019.101954.
20. Iwasaki I., Tokunaga M. Technology transfer and spillovers from FDI in transition economies: A meta-analysis // *Journal of Comparative Economics*. 2016. Vol. 44. No 4. P. 1086–1114. DOI: 10.1016/j.jce.2016.10.005.
21. Rahayu R., Day J. Determinant factors of e-commerce adoption by SMEs in developing country: evidence from Indonesia // *Procedia-Social and Behavioral Sciences*. 2015. Vol. 195. P. 142–150. DOI: 10.1016/j.sbspro.2015.06.423.
22. Belvedere V., Grando A., Papadimitriou T. The responsiveness of Italian small-to-medium sized plants: Dimensions and determinants // *International Journal of Production Research*. 2010. Vol. 48. No 21. P. 6481–6498. DOI: 10.1080/00207540903234751.
23. Zhang J., Li H., Ziegelmayer J.L. An empirical study of the relationship between IT infrastructure flexibility and IT responsiveness in SMEs: A resource-based analysis. *Proceedings of the Fourteenth Americas Conference on Information Systems*, Toronto, ON, Canada, 14–17 August 2008, pp. 1–7.
24. Cai Z., Huang Q., Liu H., Liang L. The moderating role of information technology capability in the relationship between supply chain collaboration and organizational responsiveness: Evidence from China // *International Journal of Operations and Production Management*. 2016. Vol. 36. No 10. P. 1247–1271. DOI: 10.1108/IJOPM-08-2014-0406.
25. Wade M., Hulland J. Review: the resource-based view and information systems research: review, extension, and suggestions for future research // *MIS Quarterly*. 2004. Vol. 28. No 1. P. 107–142. DOI: 10.2307/25148626.
26. Bharadwaj A.S., Sambamurthy V., Zmud R.W. IT capabilities: theoretical perspectives and empirical operationalization. *Proceedings of the 20th International Conference on Information Systems (ICIS 99)*, Charlotte, North Carolina, USA, 13–15 December 1999, pp. 378–385.
27. Sui S., Baum M. Internationalization strategy, firm resources and the survival of SMEs in the export market // *Journal of International Business Studies*. 2014. Vol. 45. No 7. P. 821–841. DOI: 10.1057/jibs.2014.11.
28. Akhtar M.W., Imran M. Knowledge management process, HRM practices and individual performance // *Australian Journal of Social Science Research*. 2015. Vol. 2. No 2. P. 8–15.
29. Åkerman N. Knowledge-acquisition strategies and the effects on market knowledge – profiling the internationalizing firm // *European Management Journal*. 2015. Vol. 33. No 2. P. 79–88. DOI: 10.1016/j.emj.2014.06.003.
30. Yu X., Yan J., Assimakopoulos D. Case analysis of imitative innovation in Chinese manufacturing SMEs: Products, features, barriers and competences for transition // *International Journal of Information Management*. 2015. Vol. 35. No 4. P. 520–525. DOI: 10.1016/j.ijinfomgt.2015.03.003.
31. Vaidyanathan S., Kidambi S.S. An empirical evaluation of adoption and diffusion of new ICTs for knowledge sharing in IT organizations // *International Journal of Web Portals*. 2018. Vol. 10. No 1. P. 1–14. DOI: 10.4018/IJWP.2018010101.
32. Chakravarty A., Grewal R., Sambamurthy V. Information technology competencies, organizational agility, and firm performance: Enabling and facilitating roles // *Information systems research*. 2013. Vol. 24. No 4. P. 976–997. DOI: 10.1287/isre.2013.0500.
33. Taylor P. The importance of information and communication technologies (ICTs): An integration of the extant literature on ICT adoption in small and medium enterprises // *International Journal of Economics, Commerce and Management*. 2015. Vol. 3. No 5. P. 274–290.
34. Gancarczyk M., Gancarczyk J. Proactive international strategies of cluster SMEs // *European Management Journal*. 2018. Vol. 36. No 1. P. 59–70. DOI: 10.1016/j.emj.2017.03.002.

35. Eco-innovation determinants in manufacturing SMEs: Systematic review and research directions / D.A. de Jesus Pacheco [et al.] // *Journal of Cleaner Production*. 2017. Vol. 142. No 4. P. 2277–2287. DOI: 10.1016/j.jclepro.2016.11.049.
36. Akram M.U., Chauhan C., Ghosh K., Singh A. Knowledge management, sustainable business performance and empowering leadership: A firm-level approach // *International Journal of Knowledge Management*. 2019. Vol. 15. No 2. P. 20–35. DOI: 10.4018/IJKM.2019040102.
37. Jorna R. (Ed.) *Sustainable innovation: The organisational, human and knowledge dimension*. Routledge, 2017.
38. Supply chain coordination through integration of innovation effort and advertising support / J. Song [et al.] // *Applied Mathematical Modelling*. 2017. No 49. P. 108–123. DOI: 10.1016/j.apm.2017.04.041.
39. Yee-Loong Chong A., Ooi K.B., Bao H., Lin B. Can e-business adoption be influenced by knowledge management? An empirical analysis of Malaysian SMEs // *Journal of Knowledge Management*. 2014. Vol. 18. No 1. P. 121–136. DOI: 10.1108/JKM-08-2013-0323.
40. Nidumolu S.R., Knotts G.W. The effects of customizability and reusability on perceived process and competitive performance of software firms // *MIS Quarterly*. 1998. Vol. 22. No 2, P. 105–137.
41. Pavlou P.A., El Sawy O.A. From IT leveraging competence to competitive advantage in turbulent environments: The case of new product development // *Information Systems Research*. 2006. Vol. 17. No 3. P. 198–227. DOI: 10.1287/isre.1060.0094.
42. Rahn R. *A comprehensive critique of the American College Health Association's National College Health Assessment survey instrument* (Doctoral dissertation). 2014.
43. Henson R.K. Teacher self-efficacy: Substantive implications and measurement dilemmas. Presented at the Annual Meeting of the Educational Research Exchange, College Station, TX, USA, 26 January 2001, p. 1–43.
44. Sharma B. A focus on reliability in developmental research through Cronbach's Alpha among medical, dental and paramedical professionals // *Asian Pacific Journal of Health Sciences*. 2016. Vol. 3. No 4. P. 271–278. DOI: 10.21276/apjhs.2016.3.4.43.
45. Ajayi A., Olayungbo D.O. ICT adoption in small and medium scale enterprises in Nigeria: An assessment // *International Journal of Research*. 2014. Vol. 1. No 9. P. 889–897.

Об авторе

Найеф Г. Азяби

Ph.D.;

Associate Professor, Department of Management Information Systems, College of Business Administration, Jazan University, Jazan 45142, Saudi Arabia;

E-mail: nazyabi@jazanu.edu.sa

ORCID: 0000-0003-0798-8114