

# Системные основания естественной классификации

Т. Л. Качанова<sup>1</sup>, К. А. Туральчук<sup>2</sup>, Б. Ф. Фомин<sup>3</sup>, О. Б. Фомин<sup>4</sup>

Санкт-Петербургский государственный электротехнический университет

«ЛЭТИ» им. В.И. Ульянова (Ленина)

<sup>1</sup>kachanova-tamara@mail.ru, <sup>2</sup>kost.tk@mail.ru, <sup>3</sup>bfomin@mail.ru, <sup>4</sup>wrk@mail.ru

**Аннотация.** Проблема естественной классификации много лет является актуальной проблемой классической, научной и теоретической систематики. Попытки создать рациональный воспроизводимый научный метод решения этой проблемы до последнего времени не приводили к успеху. Новые возможности возникли с созданием физики открытых систем, ответившей на вызовы естественной классификации разработкой онтологических предпосылок и операционального определения таксона. В статье дан обзор результатов, позволивших преодолеть сложность проблемы и предложить рациональный воспроизводимый доказанный метод решения общесистемной задачи естественной классификации на базе многомерной знание-центрической аналитики физики открытых систем. Метод рассчитан на применение в разных предметных областях к открытым природным, общественным, антропогенным, кибер-физическим и сложным техническим системам с сотнями и тысячами показателей, исходно представленным в своих естественных масштабах и реальной сложности большими массивами полимодальных гетерогенных эмпирических данных.

**Ключевые слова:** физика открытых систем; общая онтология систем; естественная классификация; онтологические предпосылки; операциональное определение; рациональный метод естественной классификации

## I. ВВЕДЕНИЕ (ВЫЗОВЫ)

**Классическая систематика.** В течение многих лет классическая систематика стремилась создать «единственно истинную» таксономическую теорию, отражающую происхождение вещей в Природе. Такая теория должна была стать основой единственно верной и окончательной «естественной классификации» (ЕК), способной раскрыть генезис вещей в Природе, объяснить их группировки, отражающие то, «что есть в природе на самом деле», охватить при этом все явления в их взаимосвязи и взаимной обусловленности. [1, 2]. Несколько столетий ЕК используется как научное понятие, общепринятый термин и научная идея, направляющая действия классификаторов, главная цель которых – овладеть процессом построения ЕК с помощью рациональных средств. [3]

**Научная систематика.** Проблема ЕК была принята научной систематикой как наиболее актуальная. Исходя из самых общих представлений, научная систематика, руководимая рациональной идеей, пыталась: найти «онтологические основания» (раскрыть структуры и

причины разнообразия вещей в Природе); разработать «эпистемологические основания» (создать научный метод ЕК) [4]. Ее главными результатами стали общая логика классифицирования (современная «классиология») и «численная систематика», отражающие соответственно качественный и количественный подходы к решению проблемы ЕК [1]. Классиология воссоздает идеальный образ таксономической реальности, служит основной формой представления «качественного» аспекта структуры разнообразия объектов Природы, явленного в существовании их группировок разного уровня общности. «Численная систематика» максимально реализует «идею математизма» и в наиболее ярком виде представляет рациональную эпистемологическую идею научной систематики. В рамках численной систематики возникли “data mining”, “data science”, проблема «инструментализма». Однако численная систематика, породившая разнообразные методы «искусственных классификаций», не получила самостоятельного значения в научной систематике. Она развивается как очень важное «методологическое приложение» к другим теориям [1].

**Естественная систематика.** Признанным направлением научной систематики является естественная систематика, основанная на онтологической и эпистемологической рациональных идеях. Ее сверхзадача – разработка всеобщей «Естественной Системы», отражающей все существенные аспекты организации порядка вещей в Природе и сходственных отношений между вещами. «Теоретическая систематика» считает понятие «Система» особо важным, полагая, что в этом понятии закон упорядоченности самой Природы выражен в наибольшей степени, и что ЕК является наилучшим описанием этого закона. [1].

Проблема ЕК имеет фундаментальное значение. Ранее ее решение более всего интересовало биологию [5]. В настоящее время ЕК является актуальной проблемой многих наук. Однако попытки создать рациональный воспроизводимый научный метод ЕК до последнего времени не привели к успеху. Онтологические предпосылки и операциональное определение ЕК, такие, чтобы на их основе можно было строить научный метод, все еще не разработаны.

## II. ФИЗИКА ОТКРЫТЫХ СИСТЕМ

Открытые системы обмениваются с внешней средой веществом, энергией, информацией. Природные, общественные, антропогенные, кибер-физические системы

рассматриваются как открытые системы. Они обладают фундаментальной сложностью. Задачей ФОС является познание, научное понимание и рациональное объяснение онтологии открытых систем с сотнями и тысячами показателей, исходно заданных эмпирическими описаниями, полученными из больших массивов слабоструктурированных полимодальных гетерогенных эмпирических данных [6, 7].

ФОС создала научный метод получения достоверного знания об онтологии систем [8–10]. ФОС раскрыла закономерности устройства, многообразие и взаимозависимость структур, множественные внутрисистемные взаимодействия, отвечающие за системогенез. Теоретические модели ФОС носят всеобщий характер, производятся из эмпирических описаний систем без обращений к экспертному знанию, субъективному анализу и интерпретациям, отвечают аксиомам систем, имеют в своей основе системообразующие симметрии и закономерности [11]. ФОС характеризует каждый показатель системы именем, значениями, системными атрибутами. Главная функция показателя – функция различения. Каждый эталон состояния системы ФОС представляет структурированным набором показателей с определенной морфологией. Эталон выражает состояние одного уникального собственного качества системы как части целого и всего целого в условиях этой части. В каждом эталонном состоянии системы гармонизированы уровни значений всех показателей, детерминирующих собственное качество системы, и раскрыт смысл этого качества в данном конкретном состоянии системы. Главная функция эталона – номинативная (именование состояния определенного уникального собственного качества системы).

Каждое актуальное состояние системы ФОС представляет моделью актуального состояния (реконструкцией состояния). Всякая реконструкция представляет систему как единое целое, порожденное уникальной «сборкой» всех эталонов состояния собственных качеств системы (далее кратко – эталоны), определяющей внутрисистемные механизмы, детерминирующие данное состояние. Значение каждого показателя в реконструкции состояния моделируется определенным набором эталонов, формирующих изменчивость данного показателя. Значения показателей кодируются уровнями величин на специальной шкале уровней (эмерджентное проявление определенного свойства в данном состоянии). Эталоны состояния и реконструкции состояния – база решения проблемы ЕК. ФОС открыла новые возможности для решения проблемы ЕК, основой которых стали:

- научный метод ФОС, раскрывающий онтологию открытых систем;
- разнообразие классифицируемых объектов, в основе которого лежит смысловое единство этих объектов на онтологическом уровне;
- научный метод ЕК, выстраивающий классификационную систему на базе знания онтологии классов.

ФОС исследует ЕК как структурную, сущностную, установленную самой Природой, интенциональную, привилегированную классологию, способную: объективно отражать смыслы классов; обеспечивать научную правильность, полноту и завершенность общих утверждений о каждом классе; раскрывать сущность классифицируемых объектов и с высокой надежностью научно понимать и рационально объяснять их свойства. ФОС ответила на вызовы ЕК разработкой онтологических предпосылок и операционального определения таксона. ФОС воспринимает таксон как онтологическое (таксономический континуум) и как типологическое единство (единство по архетипу) вещей, имеющих к нему отношение. ФОС определяет таксон интенционально (указание свойств), остенсивно (указание репрезентанта), и экстенционально (указание состава). ФОС унифицировала понятие ЕК, концептуализировала ее идею, создала систему базовых понятий (модель метода), включила задачу разработки и совершенствования доказанного воспроизводимого рационального метода ЕК в состав актуальных общесистемных задач, решаемых многомерной знание-центрической аналитикой ФОС.

### III. КОНЦЕПТУАЛИЗАЦИЯ ЕК

Интенционал предиката есть экспликация того, что интуитивно понимается под смыслом [12]. Интенциональный подход расширяет классификационное поле (КП) до классификационного универсума, состоящего не только из наличных, но и из всех мыслимых объектов. КП исходно задано именами и референтами классов. Идея онтологии КП представлена семантическим треугольником с вершинами: знак (имена классов); значение знака (референты классов); смысл (содержание классов). Основанием для ЕК является знание онтологии «Системы КП». Идея «Системы КП» представлена семантическим треугольником с вершинами: знак (имя системы); значение знака (множество носителей системы); смысл (знание онтологии системы). Идея ЕК выражена через отношения этих семантических треугольников, где (рис. 1):

- «тождество» утверждает идентичность референтов КП и носителей «Системы КП»;
- «импликация» полагает, что «Система КП» проявлена в КП;
- «экспликация» допускает возможность перехода качеств «Системы КП» на свойства классов КП и возможность развития номинативной функции языка систем (разработан в ФОС) до способности определять классы КП.

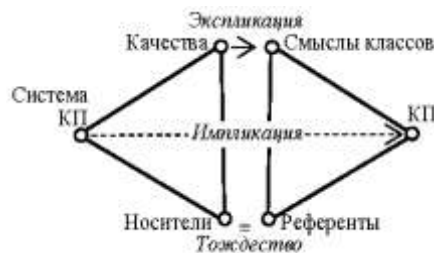


Рис. 1. Концепция ЕК

Онтологическими предпосылками ЕК являются элементы знания онтологии «Системы КП», полученного ФОС: «системные градации величин» (уровни значений первичных признаков референтов КП); «ядра эталонов», «носительство качеств» (реконструкции «системных градаций величин»).

Исследование КП как системы следует принципу двойственности:

- семантический аспект: онтология «Системы КП» установлена и эксплицирована в онтологию архетипа КП через свойства классов, типы и структуры отношений между свойствами;
- денотативный аспект: архетип КП эксплицирован в идеальные представители (прототипы) классов, на основе которых устанавливаются денотаты классов.

КП раскрыт в трех планах:

- семантическом (интенционалы классов, интенционал КП);
- денотативном (экстенционалы классов, экстенционал КП);
- оценочном (качество классифицирования).

#### IV. МОДЕЛЬ МЕТОДА ЕК

Модель основана на концепции, онтологических предпосылках и операциональном определении ЕК. Предпосылки получены из элементов знания онтологии «Системы КП». Операциональное определение разработано в опоре на онтологические предпосылки, идею использования в качестве гомологов ядер эталонов «Системы КП» (далее – «ядерности»), и еще две идеи – идею прототипа и идею экстенционала класса.

Гомологизация – ключевая идея в определении партон (вторичных признаков) таксона на основе деления понятий и отношений «общее – частное» («часть – целое», «вид – род»).

Партон – атрибут таксона, таксономический признак, средство *интенционального* определения таксона [4].

Гомологизация использует уникальные свойства ядерностей:

- гомологичность: смыслы ядерностей распространяются на инвариантные части форм их воплощений в носителях-референтах;
- референтная соотнесенность: каждый таксон имеет свои носители смысла;
- совместимость: ядерности чисто проявлены на группах референтов;
- частотность: ядерность – характерная мера таксона;
- доминантность: ядерности различают таксоны, выявляют через модальности их интенциональные признаки.

Идея прототипа выражает возможность *остенсивного* определения каждого таксона. Идея состава реализует *экстенциональный* аспект операционального определения, постулируя принцип сходности прототипа и референта.

Модель метода – система логически и семантически связанных понятий. ЕК оперирует понятиями, но не референтами или их группами.

Системный смысл классов развернут в модели по категориям «Представление классов» и «Онтология классов». Смысл классов в модели имеет три формы выражения:

- «целое» – системная онтология классов;
- «частное» – неделимые смысловые единицы онтологии классов;
- «конкретное» – экстенционалы классов.

Онтология классов представлена этапами смыслового анализа, рис. 2:

- онтология «Системы КП» (база смыслового анализа классов);
- «интенционал» (семантический аспект КП);
- «экстенционал» (денотативный аспект КП).

| Онтология классов      | Представление классов |           |                            |
|------------------------|-----------------------|-----------|----------------------------|
|                        | Целое                 | Частное   | Конкретное                 |
| Онтология «Системы КП» | Носительство качеств  | Ядерность | Системная градация величин |
| Интенционал            | Архетип               | Партон    | Класс                      |
| Экстенционал           | Таксон                | Прототип  | Признак                    |

Рис. 2. Модель метода ЕК

Научно понятая и рационально объясненная онтология классов развернута в модели по строкам на трех уровнях.

#### А. Онтологические предпосылки

*Системная градация величин* – примитивы описания референтов на единой шкале измерений (прямая связь классов с первичными признаками).

*Ядерность* – база выявления инвариантов семантической компоненты КП, выражена через примитивы системной градации величин, несет уникальный интегральный и дифференциальный смысл «Системы КП», является ее неделимой смысловой единицей.

*Носительство качеств* – база формирования семантически сочетаемых сборок ядерностей. Форма представления этого понятия – корпус текстов, выполняющий номинативную функцию по отношению к классам. Каждый референт здесь представлен уникальной сборкой ядерностей.

#### В. Интенционал

*Класс* – элемент КП, имеет имя, представлен референтами. Системное знание распределено по классам. Каждый класс задан набором ядерностей и первичными признаками, доминирующими в этом классе. Класс

отвечает принципу структурирования (референтной соотнесенности смысловых моментов).

*Партон* – характерный смысловой момент класса, сборка ядерностей 1-, 2-, 3-го и т. д. рангов (ранг – количество ядер в составе партона). Виды связи ядерностей – «соединение» и «подчинение».

*Архетип* – иерархическая структура партонов (реконструкция семантической компоненты класса), имеет вертикальное и горизонтальное измерения. Вертикальное измерение определяет центр, ядро и периферию класса. Это измерение образуют отношения «включение», «наследование», «подобие». Горизонтальное измерение проявляет моно- или полиморфизм класса. Горизонтальное измерение образуют отношения «тождество», «антиномия». Архетип класса развивается в понятие «Архетип КП» (в целом) через иерархические отношения «общее – частное», «часть – целое» и «вид – род».

#### С. Экстенционал

*Признак* – свойство, объективно присущее референтам; значения признаков дискретны (системная градация величин). Признаки–маркеры, способны выделять класс в КП. Каждый маркер имеет меру для различения класса.

*Прототип* – идеализированный объект, выражающий идею архетипа класса через признаки–конституенты. Каждый признак–конституент входит в центр архетипа и служит маркером класса. Идея архетипа оформлена через ее носители, имеет статутные признаки (существенны для референтов класса) и периферические признаки (характерны для малых групп референтов класса).

*Таксон* – структура признаков, в которой выделены области: область центра (прототип таксона), область ядра (аналог прототипа), окооядерная область (близка к прототипу), области близкой и дальней периферий (слабо связана с прототипом), смежная область (принадлежность к таксону сомнительна).

#### Д. Формообразная и формализованная коммуникации

«Формообразная» коммуникация понятий «признак–партон–носительство качеств» (первая диагональ модели) возникает при довлеющей роли факта. Она объясняет в главных моментах устройство признакового и смыслового пространств «Системы КП».

«Формализованная» коммуникация понятий «системная градация величин–партон–таксон» (вторая диагональ модели) возникает при доминирующей роли знания. Она выражает в наиболее полном и завершенном виде все то, что является знанием в раскрытом смысле системы и нацеливает все понятое знание на создание инструментальной базы понимания КП как системы.

Формообразная и формализованная коммуникации имеют общий *центр понимания знания*, в качестве которого в модели выступает партон. Через него выражены все интегральные и дифференциальные семантические признаки системы.

Тройка понятий над первой диагональю «класс–системная градация величин–ядерность» устанавливает

связь групп референтов классов с ядерностями. Тройка понятий под главной диагональю «прототип–таксон–архетип» связывает интенционал и экстенционал класса через понятие прототипа, идеально выражающего идею класса. Тройка понятий над второй диагональю «ядерность–носительство качеств–архетип» выражает онтологию семантического компонента КП. Тройка понятий под второй диагональю «класс–признак–прототип» задает морфологию денотативного компонента КП.

#### V. МЕТОД ЕК

Разработка доказанного воспроизводимого рационального метода ЕК для сложных предметных областей стала возможной благодаря построению модели метода. Далее описан первый шаг на этом направлении [13–15]. Первая редакция метода создана при следующих предположениях:

- набор первичных признаков задан экспертами; он – один и тот же для всех *референтов*, полный и представительный; шкалы измерения признаков – количественные и/или качественные;
- состав классов (задан экспертами); воспринимается исключительно как условие, требующее для каждого класса научного подтверждения (опровержения); в случаях подтверждения метод гарантирует получение интенционалов, прототипов и экстенционалов классов; в случаях опровержения – метод инициирует и направляет когнитивную и аналитическую работу по научному обоснованию классов;
- в наборе первичных признаков *носителей* признак принадлежности к классу не используется;
- КП исследуется целиком, разделение на обучающую и тестовую части КП не требуется.

Модель воплощается в методе [13]:

- через трансформацию понятий модели в определения соответствующих вычислимых объектов метода;
- путем трансформации отношений между понятиями модели в алгоритмы дистрибуции ядерностей и признаков, алгоритмы синтагматического и парадигматического анализа, алгоритмы прототипирования и таксономии.

Объекты метода:

- «формализованный дескриптор»: вычислимый объект (в модели – «*партон*»); имеет ранг 1 (отвечает положениям чистоты, частотности, референтной соотнесенности); имеет ранг 2, 3 и т. д. (если из-за гетерогенности класса не отвечает положениям чистоты, частотности, референтной соотнесенности). Ограниченная сложность дескрипторов гарантирована законом Ципфа–Мандельброта;

- «структурно-смысловая модель (ССМ)»: вычислимый объект (в модели – «интенционал»); иерархическая структура, основанная на принципе делимости понятий модели и отношении «Общее-частное» на множестве формализованных дескрипторов с отношениями: тождество, антиномия, включение, наследование, подобие; имеет центр, ядро и периферию;
- «репрезентант»: вычислимый объект (в модели – «прототип»), задан набором формализованных дескрипторов, передающих свойства класса в предельном виде; является центром ССМ;
- «структура класса»: вычислимый объект (в модели – *экстенционал класса, экстенционал КП в целом*); выделяет референтные области в экстенционалах классов (центр, ядерная и окооядерная области, ближняя и дальняя периферии, смежная область); задает границы областей на основе таксономической близости референтов к репрезентантам классов;
- «оценки качества классификации»: объем, качество, ценность (правильность, полнота, завершенность) знания о смыслах и строении классов КП.

Функциональность метода:

- дистрибуция ядерностей (признаков): алгоритм порождения объектов метода, способных различать классы и доминировать в классах;
- синтагматический анализ: алгоритм определения полных наборов формализованных дескрипторов для каждого класса и построения правильных сочетаний инвариантных частей ядерностей;
- парадигматический анализ: алгоритм оформления интенционалов классов;
- прототипирование: алгоритм построения репрезентантов классов – центров референции смысла (опорных точек) классов; мера близости к опорной точке задает устройство денотата; референты всех классов получают признаки принадлежности к одной из областей денотата.

## VI. ЗАКЛЮЧЕНИЕ

Апробация первой версии метода ЕК проведена на четырех задачах [9], [14–16]:

- «ToxEffects» – задача классификации токсических эффектов острых отравлений людей фосфорорганическими веществами;
- «Secom» – задача распознавания годных и забракованных полупроводниковых изделий по данным мониторинга производственного процесса.
- «Numbers» – задача распознавания рукописных графических образов цифр 0-9;

- «Leuco» – задача классификации типов лейкемии по геномным данным от микрочипов;

Задачи «Numbers» и «Leuco» еще до апробации были решены методами «искусственной классификации». Попытки решить задачи «ToxEffects» и «Secom» в рамках искусственной классификации не привели к результату. Метод ЕК успешно решил каждую из этих задач. Доказанными преимуществами предложенного метода ЕК являются гомология и партономия.

Гомология относится к области партономии. Она рассматривается на уровне отдельно взятого объекта, действует на всех объектах, выявляет и оформляет партоны, соотносит между собой партоны, выделенные у каждого объекта, указывает на то, что они являются разными проявлениями одного атрибута некоего обобщенного объекта.

Партономия воплощает фундаментальную концепцию естественного признака как основного средства интенционального определения таксона. Партономия играет ключевую роль. Метод ЕК достаточно просто решает вопрос о выделении партонов. Независимость и дискретность первичных признаков (главная причина таксономической неопределенности) не предполагается. При построении интенционала таксона партоны образуют иерархическую структуру, полученную по правилам вывода.

Решение задачи имеет высокую ценность, если все референты классов КП принадлежат области притяжения центров этих же классов, а смежные области всех классов КП – пустые. Наличие референтов в области отчуждения денотата КП и характер распределения референтов каждого класса КП по его референтным областям проявляют степень гомогенности классов и подтверждают правильность исходной постановки классификационной задачи. Решение раскрывает смыслы классов КП, выявляет структуру денотата КП, дает каждому классу КП научное обоснование или доказанное опровержение.

Недостатки метода в его первой редакции:

- при построении интенционала использовано отношение «общее-частное»; это отношение наиболее отвечает принципу делимости понятий; развитие метода связано с дополнительным учетом отношений «часть-целое» и «род-вид»;
- метод не рассматривает эволюцию природных объектов; отсутствие этой способности принципиально ограничивает широкое применение метода в биологии;
- метод устанавливает интенционал таксона при наличии у каждого референта КП признака принадлежности к определенному классу; развитие метода предполагает возможность получать интенционалы классов без предварительного указания признака класса в описании референтов КП.

Метод реализован в технологической платформе ФОС в составе RTD-кластера решателей общесистемных задач многомерной знание-центрической системной аналитики ФОС [9], [11].

#### СПИСОК ЛИТЕРАТУРЫ

- [1] Павлинов И.Я. Номенклатура в систематике. История, теория, практика. М.: КМК Scientific Press. 2015. 439 с.
- [2] Любарский Г.Ю. Архетип, стиль и ранг в биологической систематике. М.: КМК Inc. 1996. 436 с.
- [3] Кожара В.Л. Феномен естественной классификации // Новые идеи в научной классификации: Сб. науч. трудов. 2008. Вып.5. С. 117-143.
- [4] Мейен С.В., Шрейдер Ю.А. Методологические аспекты теории классификации // Вопросы философии. 1976. № 12. С 67-79.
- [5] Розова С.С. Классификационная проблема в современной науке. Новосибирск: Наука, Сибирское отделение. 1986. 224 с.
- [6] Качанова Т.Л., Фомин Б.Ф. Основания системологии феноменального. СПб.: Изд-во СПбГЭТУ «ЛЭТИ». 1999. 180 с.
- [7] Качанова Т.Л., Фомин Б.Ф. Введение в язык систем. СПб.: Наука. 2009. 340 с.
- [8] Fomin B. Kachanova T. Cognition of ontology of Open Systems. // Procedia Computer Science Journal. 2017. vol. 103, pp. 339-346. DOI: 10.1016/j.procs.2017.01.119
- [9] Kachanova T.L., Fomin B.F., Fomin O.B. Generating Scientifically Proven Knowledge about Ontology of Open Systems. Multidimensional Knowledge-Centric System Analytics. // Ontology in Information Science. Rijeka: InTech Publ., 2018, pp. 169-204. DOI: 10.5772/intechopen.72046
- [10] Fomin B., Turalchuk K., Kachanova T., Fomin O. Scientific understanding of ontological knowledge about open systems that is automatically mined from big data. // Proc. of the 33rd International Business Information Management Association Conference, IBIMA 2019: Education Excellence and Innovation Management Through Vision 2020. Granada, 2019, pp. 8870-8876.
- [11] Fomin B.F., Kachanova T.L., Fomin O.B. Digital Twins of Open Systems. Cyber-Physical Systems and Control, CPS&C 2019: Lecture Notes in Networks and Systems. Cham: Springer Publ., 2019, vol 95, pp. 305-314. DOI:10.1007/978-3-030-34983-7\_29
- [12] Karpap R. Philosophical Foundations of Physics. An Introduction to the Philosophy of Science. New York, London, Basic Books, Inc. Publ., 1966. 300 p.
- [13] Качанова Т.Л., Фомин Б.Ф. Системная онтология классов. // Известия СПбГЭТУ «ЛЭТИ». 2015. №7. С. 25-36.
- [14] Агеев В.О., Качанова Т.Л., Фомин Б.Ф., Туральчук К.А. Естественная классификация острых отравлений фосфорорганическими веществами. // Известия СПбГЭТУ «ЛЭТИ». 2015. №8. С. 8-17.
- [15] Kachanova T., Turalchuk K., Fomin B. Class Reconstruction in the Space of Natural System Classification. // Procedia Computer Science Journal. 2019. vol. 150, pp. 140-146. DOI: 10.1016/j.procs.2019.02.027
- [16] Туральчук К.А. Метод естественной системной классификации: Автореф. дис. ... канд. техн. наук / СПбГЭТУ «ЛЭТИ». СПб., 2020. 18 с.