

УДК 001.101; 510.85

DOI: 10.18413/2518-1092-2022-7-3-0-4

Маторин С.И.¹
Гуль С.В.²**КОНЦЕПТУАЛЬНАЯ КЛАССИФИКАЦИОННАЯ МОДЕЛЬ ЗНАНИЙ
КАК СПОСОБ СОВЕРШЕНСТВОВАНИЯ УНИВЕРСАЛЬНОЙ
ДЕСЯТИЧНОЙ КЛАССИФИКАЦИИ**¹) Белгородский университет кооперации, экономики и права, ул. Садовая 116а, г. Белгород, 308023, Россия²) Белгородский государственный национальный исследовательский университет,
ул. Победы д. 85, г. Белгород, 308015, Россия*e-mail: medintseva@bsu.edu.ru***Аннотация**

В статье рассматривается актуальная задача совершенствования универсальной десятичной классификации (УДК) в связи с ее многочисленными недостатками. Приведенные в статье примеры показывают, что в УДК нарушены общеизвестные правила построения классификаций и не находят применения системные принципы построения классификационных моделей. Данные обстоятельства превращают индексирование по УДК в сложный и исключительно трудоемкий творческий, эвристический процесс, результат которого во многих случаях в принципе не может быть корректным. Для решения поставленной задачи применен системный подход к построению концептуальных классификационных моделей (ККМ), позволяющий представлять концептуальную (понятийную) модель формы и содержания знаний, которая обеспечит легкость ее восприятия, запоминания и использования, а также однозначность решения задачи выбора индекса документа по содержанию и форме информации. Представлена базовая классификация знаний и классификации знаний по содержанию и форме, на основе которых могут быть разработаны классификаторы конкретных предметных областей.

Ключевые слова: универсальная десятичная классификация; концептуальная классификационная модель; модель знаний; классификация знаний по форме; классификация знаний по содержанию

Для цитирования: Маторин С.И., Гуль С.В. Концептуальная классификационная модель знаний как способ совершенствования универсальной десятичной классификации // Научный результат. Информационные технологии. – Т.7, №3, 2022. – С. 34-42. DOI: 10.18413/2518-1092-2022-7-3-0-4

Matorin S.I.¹
Gul S.V.²**A CONCEPTUAL CLASSIFICATION MODEL OF KNOWLEDGE
AS A WAY TO IMPROVE UNIVERSAL DECIMAL CLASSIFICATION**¹) Belgorod University of Cooperation, Economy and Law, 116a Sadovaja St., Belgorod, 308023, Russia²) Belgorod State National Research University, 85 Pobedy St., Belgorod, 308015, Russia*e-mail: medintseva@bsu.edu.ru***Abstract**

The article deals with the urgent task of improving the universal decimal classification (UDC) in connection with its numerous shortcomings. The examples given in the article show that the well-known rules for constructing classifications are violated in the UDC and the system principles for constructing classification models are not applied. These circumstances turn UDC indexing into a complex and extremely time-consuming creative, heuristic process, the result of which in many cases, in principle, cannot be correct. To solve the problem, a systematic approach to the construction of conceptual classification models (CCM) was applied, which allows representing a conceptual (conceptual) model of the form and content of knowledge, which will ensure the ease of its perception, memorization and use, as well as the uniqueness of solving the problem of choosing a document index on the content and form of information. The basic classification of knowledge and the classification of knowledge by content and form are presented, on the basis of which classifiers of specific subject areas can be developed.

Keywords: universal decimal classification; conceptual classification model; knowledge model; classification of knowledge by form; classification of knowledge by content

For citation: Matorin S.I., Gul S.V. A conceptual classification model of knowledge as a way to improve universal decimal classification // Research result. Information technologies. – Т.7, №3, 2022. – P. 34-42. DOI: 10.18413/2518-1092-2022-7-3-0-4

ВВЕДЕНИЕ

Увеличение объемов и потоков информации требует от органов научно-технической информации, библиотек, редакций, издательств и т.д. все более интенсивного использования новых компьютерных технологий для ее обработки. Большую роль в повышении эффективности обработки информации играют классификации, главной из которых, в настоящее время, является Универсальная Десятичная Классификация (УДК).

Несмотря на определенные удобства УДК, она, к сожалению, не лишена существенных недостатков, затрудняющих поиск индекса, соответствующего содержанию индексируемого документа, и усложняющих тем самым процесс индексирования. Некоторые из них настолько очевидны, что упоминаются в самих изданиях этой классификации при ее описании, а именно [1]:

→ множественность локализации понятий, т.е. многократное расположение одних и тех же понятий по различным отраслям знания в зависимости от аспекта рассмотрения;

→ дробность детализации схемы, т.е. необходимость представления индекса документа в виде нескольких (иногда двухэтажных) десятичных дробей, имеющих различную мнемонику;

→ органическая связанность разделов (изменение одного раздела влечет за собой изменение другого), например, связь классов 5 и 6 основного ряда настолько сильна, что зачастую вызывает затруднения при выборе индекса;

→ при построении разделов основной таблицы (основного ряда) УДК, а также вспомогательных таблиц (специальных и общих определителей) в основу их создания не заложен единый системообразующий принцип, что находит свое выражение в особом дополнительном понятии «параллельное подразделение» и необходимости использования специального знака конгруэнтности « \cong »;

→ многие разделы разработаны несколько десятилетий тому назад.

Кроме того, во многих разделах и подразделах УДК, несмотря на постоянные попытки ее совершенствования, нарушены общеизвестные правила построения классификаций (правила деления понятий: соразмерность членов деления и взаимное их исключение и т.д.) [2] и тем более не находят применения системные принципы построения концептуальных классификационных моделей (ККМ) [3].

ПРОБЛЕМЫ УДК

Например, анализируя самую последнюю версию УДК [4], можно показать, что как правило, сведенными вместе без четкого различия оказываются такие характеристики знания как «форма» и «содержание», что может создавать путаницу при индексировании. Например, в общем разделе *«Наука и знание. Организация. Информационные технологии. Информация. Документация. Библиотечное дело. Учреждения. Публикации»* (0), в классе (00) – *«Общие вопросы науки и культуры. Пропедевтика»* – представлены как формы представления знаний: *«Документация Научно-техническая информация (НТИ). Печать в целом. Авторство»* (002), *«Системы письма и письменности. Знаки и символы. Семиотика в целом. Коды. Графическое представление мысли»* (003); так и содержательные аспекты *«Информационные технологии. Вычислительная техника. Обработка данных»* (004), в том числе *«Искусственный интеллект»* (004.8), и *«Деятельность и организация. Общая теория связи и управления (кибернетика)»* (007), в том числе *«Изменчивость физической конфигурации и состояния. Изменение состояния»* (007.2).

В подразделе «*Организация умственного труда*» (001.81) на одном уровне деления находятся такие пункты как «*Подбор справочных источников. Изучение содержания, аннотаций, выдержек, резюме, первых набросков, литературных черновиков*» (001.814), «*Подготовка докладов. Оформление докладов*» (001.817), «*Изложение результатов работы. Составление отчетов. Техника отчетности*» (001.818); так и «*Средства для работы*» (001.812) и «*Собирание газетных и прочих вырезок*» (001.813), которые не являются видами умственного труда. При этом классифицирование как один из важных способов организации умственного труда отсутствует.

Не удивительно, что при этом в совершенно разных классах и на разных уровнях оказываются понятия одного содержательного поля, например, «*культура*» (в классе 008), «*искусство*» (в классе 7), а «*музеи*» (в классе 06); «*письменность*» (в классе 003), а «*языкознание*» (в классе 8).

При этом в некоторых классах объединенными оказываются совершенно не связанные между собой понятия. Например, в классе (06) объединены «*музеи*» с «*финансовыми комитетами*»; в классе (65) спокойно соседствуют на одном уровне «*Телекоммуникация и дистанционное управление*» (654) и «*Бухгалтерия. Бухгалтерский учет*» (657). А классы (302.42) и (302.44) называются одинаково «*Виды исследований*».

Судя по этой классификации, «*география*» (9) оказывается не «*естественной наукой*» (5). При этом в последний класс включены разделы, явно связанные с географией: «*Природа. Охрана природных ресурсов. Угрозы окружающей среды и защита от них*» (502/504), «*Астрономия. Астрофизика. Исследование космического пространства. Геодезия*» (52), «*Науки о Земле. Геологические науки*» (55). Таким же образом «*история*» (9) является и не «*естественной*» (5), и не «*общественной*» (3) наукой.

Кроме того, «*телекоммуникации*», представляющие собой информационные технологии, попали в класс (65), остальные «*информационные технологии*» в класс (004). В классе «*Документация. Научно-техническая информация (НТИ). Печать в целом. Авторство*» (002) нет подкласса, касающегося собственно НТИ, только подкласс «*Центры документации. Органы НТИ*» (002.6).

Большую трудность представляет индексирование в соответствии с УДК документов по междисциплинарным исследованиям или интегрированным научным дисциплинам. При этом добавление новых научных дисциплин в УДК является сложной и не имеющей однозначного решения задачей. Например, в соответствии с действующей УДК «*Системный анализ*» запрятан в подклассе (303.372), т.е. должен использоваться только в рамках общественных наук, хотя давно известно его применение и в технике, и в информационных технологиях.

На основании сказанного выше можно утверждать, что с помощью УДК невозможно объективизировать процесс определения индекса документа. Индексирование по УДК это сложный и исключительно трудоемкий творческий, эвристический процесс, результат которого во многих случаях в принципе не может быть корректным.

Невозможность алгоритмизации индексирования по УДК исключает автоматизацию этого процесса и, таким образом, препятствует широкому внедрению компьютерных технологий, необходимых для повышения эффективности обработки информации.

ПЕРСПЕКТИВЫ СОВЕРШЕНСТВОВАНИЯ УДК С ПОМОЩЬЮ ККМ

В настоящее время, в связи с успехами в области системных исследований, а также теории классификации и представления знаний существует возможность создания новой классификации знаний, учитывающей критерии и системные закономерности ККМ, в рамках которой отмеченные недостатки могут быть устранены.

Системная ККМ будет обладать следующими преимуществами и возможностями:

→ представлять собой концептуальную (понятийную) модель формы и содержания знаний, что обеспечит легкость ее восприятия, запоминания и использования;

→ обеспечивать однозначность решения задачи выбора индекса документа по содержанию и форме информации;

→ автоматизировать процесс индексирования документов по ключевым словам или понятиям (и не только документов, а и всех других форм, в которых выражаются человеческие знания: картин, музыкальных произведений, любых изделий);

→ обеспечивать простое и, при необходимости, оперативное введение новых отраслей знаний, изменение существующих отраслей (за счет простых операций абстрагирования (обобщения) и конкретизации (детализации) над понятиями данной предметной области) без изменения других областей классификации;

→ возможность использования отдельных разделов и подразделов ККМ в качестве концептуальных моделей конкретных предметных областей в базах знаний современных знаниеориентированных информационных систем и технологий;

→ возможность построения информационных поисковых систем и технологий по единому принципу и основанию, например, в Internet, что обеспечит значительное повышение эффективности хранения и поиска информации в международном масштабе.

Рассмотрим пути построения системной ККМ знаний.

БАЗОВАЯ КЛАССИФИКАЦИЯ ЗНАНИЙ

При построении такой классификации (при выработке принципа ее построения) необходимо исходить из роли, места, функции (функциональных свойств) знаний в природе. Знаниям, как правило, приписывается двойственная функция, с одной стороны, содержательная, связанная с отражением действительности в сознании человека, а, с другой стороны, формальная, связанная с некоторым способом их выражения (для фиксации и последующей передачи).

В соответствии с системным подходом для получения требуемой ККМ объектов некоторой предметной области, обязательно необходимо учитывать два вида их свойств. Во-первых, это, так называемые, качественные свойства (различные качества) и, во-вторых, это, так называемые, граничные свойства (структура, функция, количество и т.д.). Легко видеть, что свойство отражения (содержательное) для знания является качественным свойством, а свойство выражения (формальное) – граничным. Таким образом, в абстрактном классе «ЗНАНИЕ» должны быть, в первую очередь, выделены класс «СОДЕРЖАНИЕ ЗНАНИЙ», обеспечивающий выполнение функции отражения, и класс «ФОРМА ЗНАНИЙ», обеспечивающий функцию выражения. Тогда, в соответствии с параметричностью системной ККМ, классификация видов формы знаний должна соответствовать классификации видов выражения этих знаний, а классификация видов содержания знаний должна соответствовать классификации видов отражений (см. рис. 1).

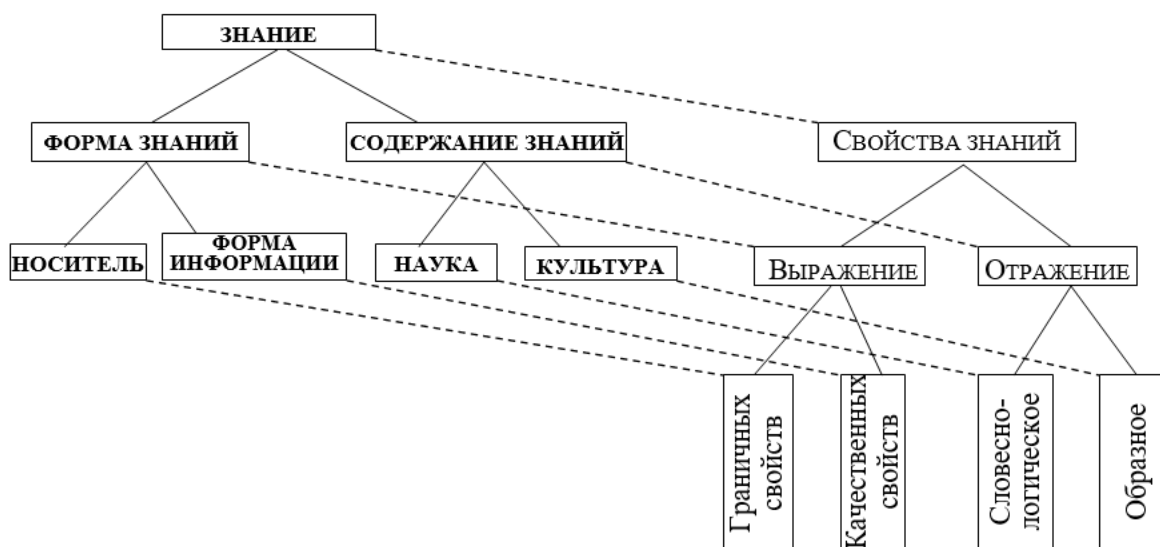


Рис. 1. Базовая классификация знаний
Fig. 1. Basic classification of knowledge

В качестве основных видов «**ФОРМЫ ЗНАНИЙ**» следует рассматривать виды «**НОСИТЕЛЕЙ**» знаний, соответствующие выражению граничных свойств формы знаний, и виды «**ФОРМЫ ИНФОРМАЦИИ**», соответствующие выражению качественных свойства формы знаний. В качестве основных видов «**СОДЕРЖАНИЯ ЗНАНИЙ**» следует рассматривать «**НАУКУ**», обеспечивающую абстрактное словесно-логическое отражение действительности и соответствующую, таким образом, граничным свойствам содержания знаний, и «**КУЛЬТУРУ**», обеспечивающую в большей степени чувственное образное отражение и соответствующую, таким образом, качественным свойствам содержания знаний.

Дальнейшее классифицирование формы знаний (носителей знаний и формы информации) и содержания знаний (науки и культуры) должно осуществляться по тому же принципу путем деления каждого вида выражения знаний и каждого вида отражения знаний на граничные (формальные) и качественные (содержательные) свойства.

КЛАССИФИКАЦИЯ ЗНАНИЙ ПО СОДЕРЖАНИЮ

Рассмотрим классификацию содержания знаний по видам отражения (см. рис. 2).

В связи с наличием двух видов отражения – словесно-логического и образного – все содержание знаний делится на «**НАУКУ**», как знания, обеспечивающие собственно словесно-логическое отражение действительности, и «**КУЛЬТУРУ**», как знания, обеспечивающие преимущественно образное, чувственное отражение этой действительности.

В зависимости от степени абстрактности словесно-логического отражения науку следует разделить на «**ТЕОРЕТИЧЕСКУЮ**», как науку, обеспечивающую большую степень абстрактности (формальности), и «**ЭМПИРИЧЕСКУЮ**» – имеющую большую конкретность (содержательность) отражения. В зависимости от степени рефлексивности образного отражения, соответствующей степени абстрактности словесно-логического отражения, культуру следует разделить на «**ИСКУССТВО**», обеспечивающее отражение с большей степенью рефлексии, и «**РЕЛИГИЮ**», обеспечивающую отражение с меньшей степенью рефлексии.

Далее аналогично, теоретическую науку в зависимости от обеспечиваемой степени абстрактности отражения следует разделить на «**ЛОГИКО-МАТЕМАТИЧЕСКУЮ**» (логические исчисления, различные сугубо математические дисциплины и т.д.), как науку, обеспечивающую чисто формальное отражение, и «**ПРЕДМЕТНО-ТЕОРЕТИЧЕСКУЮ**» (теоретическая физика, космология и т.д.) – обеспечивающую формально-содержательное отражение. Эмпирическую науку следует разделить на науку «**ТЕХНИЧЕСКУЮ**», являющуюся более абстрактной и «**ЕСТЕСТВЕННУЮ**» – являющуюся более конкретной. По тому же принципу техническую науку следует разделить на «**СОБСТВЕННО ТЕХНИЧЕСКУЮ**» (более формальную, например, радиоэлектроника) и «**ИНФОРМАЦИОННУЮ**» (менее формальную, например, искусственный интеллект), а естественную науку – на более формализованную «**ПРИРОДНУЮ**» науку (о природе: физика; химия; физическая география; биология, зоология и т.д.) и менее формализованную «**СОЦИАЛЬНО-ГУМАНИТАРНУЮ**» (общественные науки, в частности экономическая география; история; языкознание; психология, педагогика и т.д.).

В соответствии со степенью рефлексивности искусство следует разделить на «**СЛОВЕСНО-РЕЧЕВОЕ**» (литература; театр; эстрада и т.д.), связанное с левополушарной фиксацией результатов отражения, и «**МУЗЫКАЛЬНО-ИЗОБРАЗИТЕЛЬНОЕ**», связанное с правополушарной фиксацией отражения.

Религию, по тому же принципу, на религию «**ФИЛОСОФСКУЮ**», которая включает в себя развитые философские системы, обеспечивающие более сильную рефлексивность (например, дао), и «**ТЕОЛОГИЧЕСКУЮ**», представляющую собой верования, обходящиеся более слабой степенью рефлексии или вообще без таковой (например, христианство).

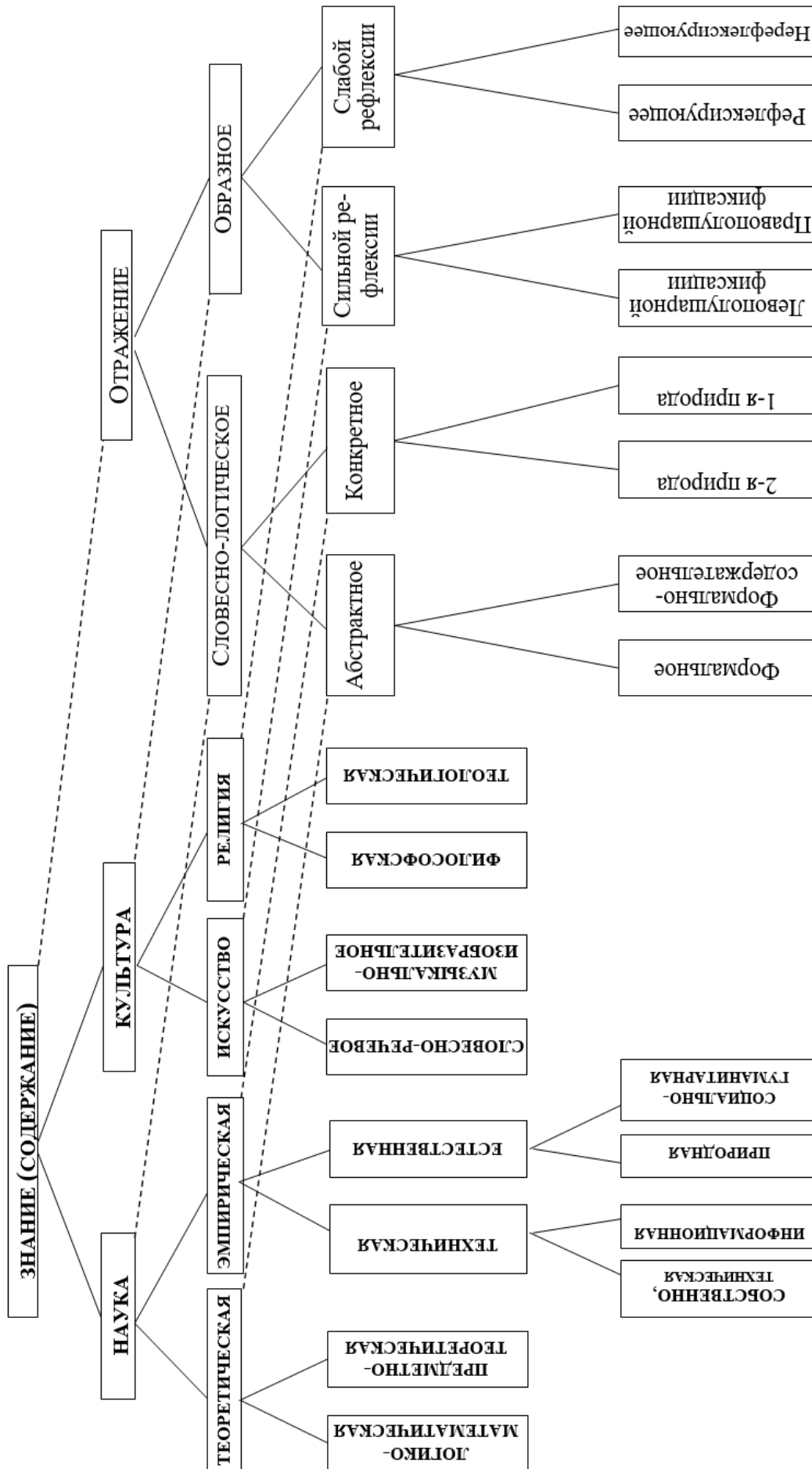


Рис. 2. Классификация содержания знаний по видам отражения
Fig. 2. Classification of knowledge content by types of reflection

Таким образом, содержание знания можно упорядочить по единому принципу и расположить все области знаний последовательно от самых абстрактных формализованных логико-математических научных дисциплин до совершенно чувственных и неререфлективных религиозных верований.

КЛАССИФИКАЦИЯ ЗНАНИЙ ПО ФОРМЕ

Рассмотрим классификацию формы знаний по видам выражения знаний (см. рис. 3).

В связи с выделением граничных и качественных свойств в рамках функциональной способности знаний выражаться в некоторой форме, форма знаний разделена на «НОСИТЕЛЕЙ» знаний, выражающих граничные свойства формы знаний, и «ФОРМУ ИНФОРМАЦИИ», выражающую качественные свойства формы знаний.

В зависимости от жесткости (материальности) выражаемых граничных свойств следует разделить носителей на «ВЕЩЕСТВЕННЫХ», обладающих более жесткими материальными граничными свойствами, и на «ПОЛЕВЫХ», обладающих менее жесткими энергетическими граничными свойствами. Последние объединяют «РЕАЛЬНЫХ» фиксированных носителей (дискеты, диски, флешки и т.д.) и «ВИРТУАЛЬНЫХ» распределенных (например, адрес www). В зависимости от степени формальности выражаемых качественных свойств форму информации следует разделить на «СПОСОБ», как вид формы информации, соответствующий выражению структурных (граничных, формальных) свойств знаний (в рамках данного класса будут различаться знаковые и не знаковые способы выражения информации и, в первую очередь, такие понятия как: «ЯЗЫКОВЫЙ» и «АУДИОВИЗУАЛЬНЫЙ»), и «СТИЛЬ», как вид формы информации, обеспечивающий выражение стилистических (менее формальных, качественных) характеристик знаний (в рамках данного класса будут различаться «ФОРМАЛЬНЫЕ» и «СОДЕРЖАТЕЛЬНЫЕ» стили).

В рамках вещественных (материальных) носителей следует рассматривать – «ПРЕДМЕТЫ» (искусственные изделия, в том числе документы, как более формальные) и – «ОБЪЕКТЫ» (естественные неорганические и органические, как менее формальные). В рамках языкового способа выражения информации следует рассматривать «ИСКУССТВЕННО ЯЗЫКОВЫЕ» способы, как более формальные, и «ЕСТЕСТВЕННОЯЗЫКОВЫЕ», как менее формальные более содержательные.

ЗАКЛЮЧЕНИЕ

Предлагаемый системологический метод систематизации знаний и базовая классификация знаний, отлаженная и доведенная, конечно, до конкретных терминов соответствующих дисциплин и предметных областей, обеспечит объективизацию и алгоритмизацию индексирования документов по их понятийному содержанию, что позволит действительно автоматизировать процесс индексирования по ключевым словам (терминам).

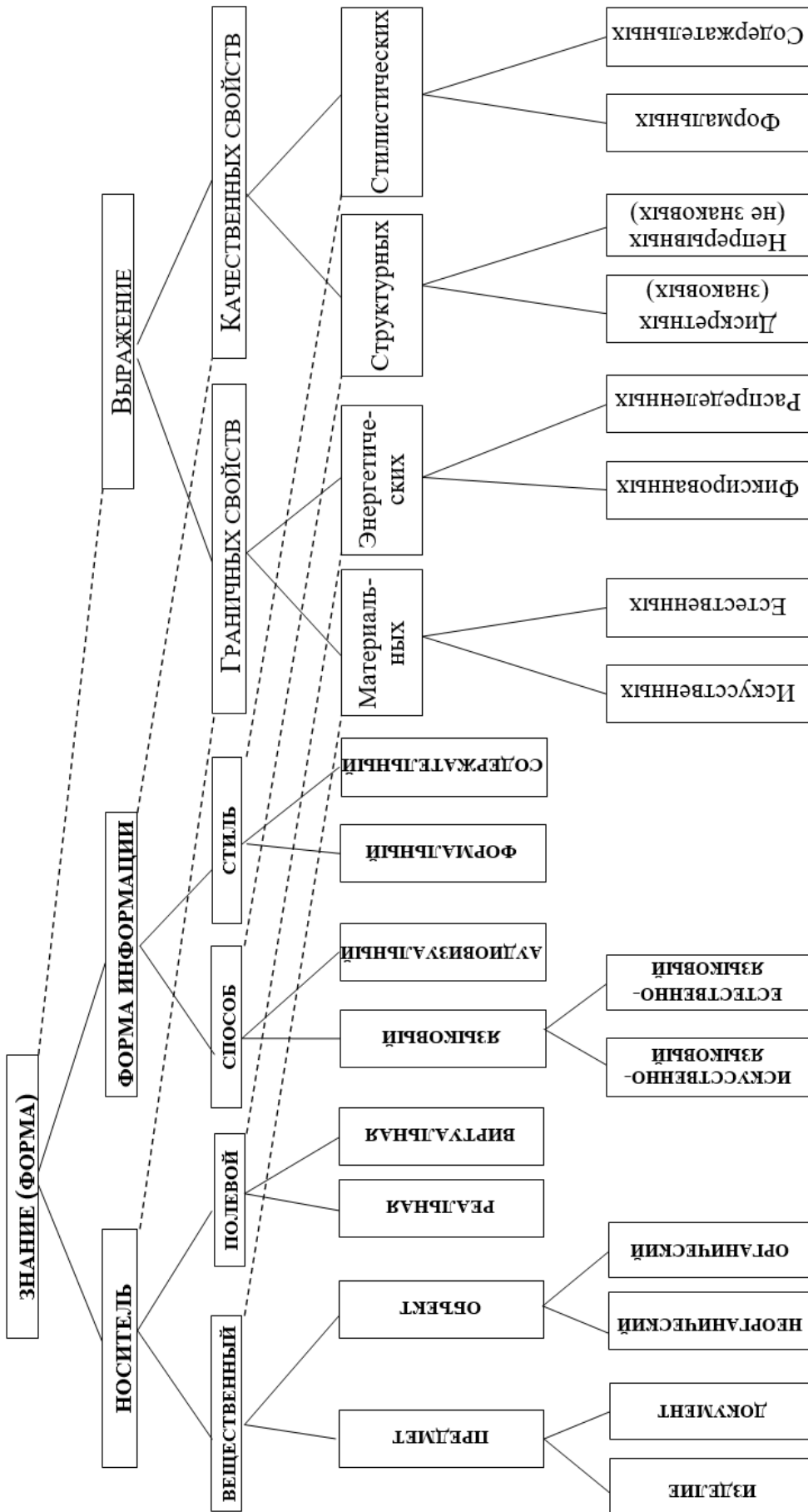


Рис. 3. Классификация формы знаний по видам выражения
Fig. 3. Classification of the form of knowledge by types of expression

БЛАГОДАРНОСТИ

Исследования выполнены при финансовой поддержке проекта Российского фонда фундаментальных исследований № 19-29-01047мк.

Список литературы

1. Универсальная Десятичная Классификация/ Третье советское издание. Вып.2, ч.1. М.: Изд-во стандартов, 1983. 248с.
2. Кондаков Н.И., Горский Д.П. Логический словарь. – М.: Наука, 1975. – 656 с.
3. Гуль С.В., Маторин С.И. Проблемы концептуального классификационного моделирования // Материалы IX международной конференции «Информационно-аналитические системы и технологии». 14-18 марта 2021 года, Белгород: Издательство БУКЭП, 2022. С. 49-57.
4. УДК Классификатор 2021 – Издательство ТРИУМФ. URL: <https://www.triumph.ru/html/serv/udk.html>

References

1. Universal Decimal Classification / Third Soviet Edition. Issue 2, part 1. M.: Publishing House of Standards, 1983. 248s.
2. Kondakov N.I., Gorsky D.P. Logical Dictionary. – M.: Nauka, 1975. – 656 p.
3. Gul S.V., Matorin S.I. Problems of conceptual classification modeling // Materials of the IX International Conference "Information and analytical systems and Technologies". March 14-18, 2021, Belgorod: BUKER Publishing House, 2022. pp. 49-57.
4. UDC Classifier 2021 - TRIUMPH Publishing House. URL: <https://www.triumph.ru/html/serv/udk.html>

Маторин Сергей Игоревич, д.т.н., профессор, профессор кафедры информационных систем и технологий
Гуль Светлана Владимировна, старший преподаватель кафедры информационных и робототехнических систем

Matorin Sergey Igorevich, Doctor of Technical Sciences, Professor, Professor of the Department of Information Systems and Technologies

Gul Svetlana Vladimirovna, Senior Lecturer, Department of Information and Robotic Systems