МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РФ

ФГБОУ ВО «Кубанский государственный агарный университет имени И. Т. Трубилина»

Е. В. Луценко

АВТОМАТИЗИРОВАННЫЙ СИСТЕМНО-КОГНИТИВНЫЙ АНАЛИЗ В ВЕТЕРИНАРИИ

Монография

Краснодар КубГАУ 2023 УДК 004.8 (075.8) ББК 32.965 Л86

Рецензенты:

- С. А. Курносов профессор кафедры омпьютерных технологий и систем Кубанского государственного аграрного университета, канд. экон. наук, профессор;
- Г. А. Аршинов профессор кафедры омпьютерных технологий и систем Кубанского государственного аграрного университета, д-р техн. наук, профессор

Луценко Е. В.

Л86 Автоматизированный системно-когнитивный анализ в ветеринарии : монография / Е. В. Луценко — Краснодар : КубГАУ, 2023. — 160 с.

ISBN 978-5-907757-23-3

Рассмотрен алгоритм применения Автоматизированного системно-когнитивного анализа и его программного инструментария интеллектуальной системы «Эйдос» для решения разнообразных задач области ветеринарии: идентификации, прогнозирования, принятия решений исследования объекта И моделирования путем исследования его модели.

Издание предназначено для интересующихся вопросами применения когнитивных технологий и других методов искусственного интеллекта.

УДК 004.8:34(075.8)

- © Луценко E. B., 2023
- © ФГБОУ ВО «Кубанский государственный аграрный университет имени И. Т. Трубилина, 2023

ISBN 978-5-907757-23-3 DOI 10.13140/RG.2.2.28032.92163

ОГЛАВЛЕНИЕ

1. ВВЕДЕНИЕ	4
1.1. Описание исследуемой предметной области	4
1.2. Объект и предмет исследования	5
1.3. ПРОБЛЕМА, РЕШАЕМАЯ В РАБОТЕ И ЕЕ АКТУАЛЬНОСТЬ	6
1.4. ЦЕЛЬ РАБОТЫ	7
2. МЕТОДЫ	7
2.1. Обоснование требований к методу решения проблемы	7
2.2. Литературный обзор методов решения проблемы, их характеристика и	1
ОЦЕНКА СТЕПЕНИ СООТВЕТСТВИЯ ОБОСНОВАННЫМ ТРЕБОВАНИЯМ	7
2.3. Автоматизированный системно-когнитивный анализ (АСК-анализ) ка	АК
МЕТОД РЕШЕНИЯ ПРОБЛЕМЫ	8
2.4. Задачи работы	36
3. РЕЗУЛЬТАТЫ	37
3.1. КОГНИТИВНАЯ ВЕТЕРИНАРИЯ - ВЕТЕРИНАРИЯ ЦИФРОВОГО ОБЩЕСТВА: ДЕФИНИ	щия
БАЗОВЫХ ПОНЯТИЙ	37
3.2. ФОРМИРОВАНИЕ СЕМАНТИЧЕСКОГО ЯДРА ВЕТЕРИНАРИИ ПУТЕМ АСК-АНАЛИЗА	1
паспортов научных специальностей ВАК РФ и автоматическая	
КЛАССИФИКАЦИЯ ТЕКСТОВ ПО НАПРАВЛЕНИЯМ НАУКИ	86
3.3. Агломеративная когнитивная кластеризация нозологических образо	ВВ
ветеринарии	129
3.4. Агломеративная когнитивная кластеризация симптомов и синдромо	ВВ
ВЕТЕРИНАРИИ	143
4. ОБСУЖДЕНИЕ	155
5. ВЫВОДЫ	156
СПИСОК ЛИТЕРАТУРЫ	157

1. ВВЕДЕНИЕ

1.1. Описание исследуемой предметной области

Данная работа является продолжением серии работ автора по применению Автоматизированного системно-когнитивного анализа (АСКанализ) и его программного инструментария — интеллектуальной системы «Эйдос» для решения широкого спектра задач в различных предметных областях, в частности в области ветеринарии [1-13].

Объектом моделирования в ветеринарии является животное. В работе ставится и решается задача выявления пичинно-следственных зависимостей между факторами, действующими на объект моделирования и результатами действия этих факторов, т.е. состояниями животного, вызванными этими факторами.

На основе знания этих зависимостей в работе ставятся и решаются разнообразные задачи в области ветеринарии: идентификации, диагностики, прогнозирования, принятия решений и исследования объекта моделирования путем исследования его модели. В монографии [16] подробно рассмотрено решение следующих задач:

- 1. Когнитивная ветеринария ветеринария цифрового общества: дефиниция базовых понятий.
- 2. Формирование семантического ядра ветеринарии путем АСКанализа паспортов научных специальностей ВАК РФ и автоматическая классификация текстов по направлениям науки.
- 3. Синтез семантических ядер научных специальностей ВАК РФ и автоматическая классификации статей по научным специальностям с применением АСК-анализа и интеллектуальной системы "Эйдос" (на примере научного журнала КубГАУ и его научных специальностей: механизации, агрономии и ветеринарии).
- 4. АСК-анализ в ветеринарии (на примере разработки диагностических тестов).
- 5. Реализация тестов и супертестов для ветеринарной и медицинской диагностики в среде системы искусственного интеллекта "Эйдос-X++" без программирования.
- 6. Агломеративная когнитивная кластеризация нозологических образов в ветеринарии.
- 7. Агломеративная когнитивная кластеризация симптомов и синдромов в ветеринарии.
 - 8. АСК-анализ и классификация пород крупного рогатого скота.
 - 9. АСК-анализ антибиотиков в ветеринарии.
- 10. АСК-анализ влияния пробиотиков в рационах на телосложение бычков.

- 11. Разработка ветеринарного теста для диагностики желудочнокишечных заболеваний лошади на основе данных репозитория UCI с применением АСК-анализа.
- 12. Математическое и численное моделирование взаимосвязи морфологического, биохимического и микроэлементного состава крови бычков герефордской породы и их размеров.
- 13. Когнитивная информационно-измерительная квалиметрическая система для определения содержания жира и белка в коровьем молоке по параметрам тензиограмм динамического поверхностного натяжения на границе раздела молоко/воздух.

Однако в данной монографии из-за ограничения на ее объем подробно рассмотрено решение лишь некоторые из этих задач. Полный вариант данной монографии размещен в РесечГейт: [16].

1.2. Объект и предмет исследования

Какое-либо направление исследований становится наукой только в том случае, если у нее есть свой объект и предмет исследования, и свой метод исследования.

Автоматизированный системно-когнитивный анализ (АСК-анализ) является универсальным методом научного исследования, т.е. может быть успешно применен для проведения исследований в самых различных предметных областях, в которых исследования проводятся с помощью естественного интеллекта. Во многом это возможно благодаря тому, что АСК-анализ имеет свой программный инструментарий, в качестве которого в настоящее время выступает интеллектуальная система «Эйдос». Система «Эйдос» является не только системой искусственного интеллекта, но и системой автоматизации научных исследований. Она разрабатывалась в такой постановке, что ее можно успешно применять для автоматизации научных исследований в самых различных предметных областях, в самых различных направлениях исследований, в частности в области ветеринарии.

Поэтому можно взять объект и предмет исследования практически любой науки и исследовать их с применением АСК-анализа и системы «Эйдос» и можно обоснованно утверждать, что в результате этого получится новое направление науки, с рем же объектом и предметом исследования, но с новым методом и инструментарием исследования. Название этого нового направления науки можно образовать из старого названия путем добавления к нему слова «когнитивный», означающего метод исследования.

Например: когнитивная психология¹, когнитивная педагогика², экономика³, когнитивная лингвистика⁴, когнитивная когнитивная

5

 $[\]frac{1}{2} \frac{\text{https://scholar.google.ru/scholar?hl=ru\&as_sdt=0\%2C5\&q=когнитивная+психология\&btnG=}}{\text{https://scholar.google.ru/scholar?hl=ru&as_sdt=0\%2C5\&q=Когнитивная+педагогика\&btnG=}}$

ветеринария $[1-13]^5$, агрономия [14], когнитивная когнитивная механизация [15], , и т.д., и т.п.

В ветеринарии:

объект исследования - заболевания животных (и птицы), прежде всего домашних.

причин выявление заболеваемости предмет исследования профилактика заболеваемости, животных (и птицы), диагностика и лечение.

Предлагается рассматривать лечение, как процесс управления состоянием животных (и птицы) путем воздействии на них лечебными факторами. Таким образом, в ветеринарии животные и птица, по сути, представляют собой объект управления.

Управление этим специфическим объектом управления включает

- получение информации о состоянии животного;
- диагностику и прогнозирование развития этого состояния;
- выработку управляющих решений по переводу животного из состояния заболевания в здоровое состояние.

1.3. Проблема, решаемая в работе и ее актуальность

Проблема определяется в науке как несоответствие фактического и желаемого, противоречие между ними.

Конкретно в ветеринарии решаемая проблема состоит в том, что фактически животные (и птица) часто являются больными, желательно, чтобы они были здоровыми.

Решение проблемы – это способ или метод и методика перевода фактической нежелательной ситуации в желательное целевое состояние.

На практике решение проблемы осуществляется путем применения переводящих фактическую нежелательную ситуацию факторов, желательное целевое состояние.

Таким образом, решение проблемы ветеринарии целесообразно искать путем выявления факторов, влияющих на фактическую ситуацию, и изучения силы и направления их влияния на перевод фактической ситуации в целевую. Именно это обеспечивает АСК-анализ и его программный инструментарий – интеллектуальная система «Эйдос».

Кроме того некоторые из факторов, рекомендуемых для перевода фактической ситуации в целевую, может быть нет возможности применить, и тогда возникают вопросы о том, что будет, если их не применить, или может быть есть возможность заменить их другими факторами, действующими на объект управления аналогично,

 $^{^{3}}$ <a href="https://scholar.google.ru/scholar?hl=ru&as_sdt=0%2C5&q=когнитивная+экономика&btnG="https://scholar.google.ru/scholar?hl=ru&as_sdt=0%2C5&q=когнитивная+лингвистика&btnG="https://scholar.google.ru/scholar?hl=ru&as_sdt=0%2C5&q=когнитивная+лингвистика&btnG="https://scholar.google.ru/scholar?hl=ru&as_sdt=0%2C5&q=когнитивная+лингвистика&btnG="https://scholar.google.ru/scholar?hl=ru&as_sdt=0%2C5&q=когнитивная+экономика&btnG="https://scholar.google.ru/scholar?hl=ru&as_sdt=0%2C5&q=когнитивная+лингвистика&btnG="https://scholar.google.ru/scholar?hl=ru&as_sdt=0%2C5&q=когнитивная+лингвистика&btnG="https://scholar.google.ru/scholar?hl=ru&as_sdt=0%2C5&q=когнитивная+лингвистика&btnG="https://scholar.google.ru/scholar?hl=ru&as_sdt=0%2C5&q=когнитивная+лингвистика&btnG="https://scholar.google.ru/scholar?hl=ru&as_sdt=0%2C5&q=когнитивная+лингвистика&btnG="https://scholar.google.ru/scholar?hl=ru&as_sdt=0%2C5&q=когнитивная+лингвистика&btnG="https://scholar.google.ru/scholar?hl=ru&as_sdt=0%2C5&q=когнитивная+лингвистика&btnG="https://scholar.google.ru/scholar.google.ru/scholar.google.ru/scholar.google.ru/scholar.google.ru/scholar.google.ru/scholar.google.ru/scholar.google.ru/scholar.google.ru/scholar.google.ru/scholar.google.ru/scholar.google.

⁵ https://<u>scholar.google.ru/scholar?hl=ru&as_sdt=0%2C5&q=когнитивная+ветеринария&btnG</u>=

имеющимися в нашем распоряжении. Это обеспечивает развитый алгоритм принятия решений АСК-анализа, реализованный в интеллектуальной системе «Эйдос».

1.4. Цель работы

Целью работы является решение поставленной проблемы.

Достижение поставленной цели обеспечивается решением ряда задач и подзадач, которые являются этапами достижения цели. Конкретная формулировка этих задач зависит от метода решения проблемы, поэтому обоснованно мы сформулируем их в конце раздела, т.е. после обоснованного выбора и описания этого метода решения проблемы.

2. МЕТОДЫ

2.1. Обоснование требований к методу решения проблемы

Из специфики поставленной проблемы и необходимости обеспечения сопоставимости обработки в одной модели эмпирических исходных данных, представленных в разных типах шкал числовых и текстовых (лингвистических) и в разных единицах измерения, вытекают следующие *требования* к методу решения проблемы:

- 1. Метод должен обеспечивать устойчивое выявление в сопоставимой форме силы и направления причинно-следственных зависимостей в неполных зашумленных (неточных) взаимозависимых (нелинейных) данных очень большой размерности числовой и не числовой природы, измеряемых в различных типах шкал (номинальных, порядковых и числовых) и в различных единицах измерения.
- 2. Метод решения проблемы не должен предъявлять жестких требований к исходным данным, которые невозможно выполнить, а должен обеспечивать обработку тех данных, которые реально есть.
- 3. Метод должен реально на практике решать поставленную проблему, а значит, он должен иметь поддерживающий его программный инструментарий, находящийся в полном открытом бесплатном доступе.

2.2. Литературный обзор методов решения проблемы, их характеристика и оценка степени соответствия обоснованным требованиям

Поиск в Internet программных систем, *одновременно* удовлетворяющих ранее обоснованным требованиям показал, что альтернатив Автоматизированному системно-когнитивному анализу и его программному инструментарию — системе «Эйдос» в настоящее время здесь нет [17, 18].

2.3. Автоматизированный системно-когнитивный анализ (АСК-анализ) как метод решения проблемы

2.3.1. Кратко о методе АСК-анализа

Автоматизированный системно-когнитивный анализ (АСК-анализ) предложен $npo\phi$. E.B. Луценко в 2002 году в ряде статей 1997-2001 годов фундаментальной монографии [19].

Сам термин: «Автоматизированный системно-когнитивный анализ (АСК-анализ)» был предложен проф.Е.В.Луценко. На тот момент он вообще не встречался в Internet. Сегодня по соответствующему запросу в Яндексе находится 9 миллионов сайтов с этим сочетанием слов 7 .

АСК-анализ включает:

- теоретические основы, в частности базовую формализуемую когнитивную концепцию;
- математическую модель, основанную на системном обобщении теории информации (СТИ);
- методику численных расчетов (структуры баз данных и алгоритмы их обработки);
- программный инструментарий, в качестве которого в настоящее время выступает универсальная когнитивная аналитическая система «Эйдос» (интеллектуальная система «Эйдос»).

Более подробно математический метод АСК-анализа описан в работах [19, 20] и ряде других [17, 18]. Около половины из 675 опубликованных автором научных работ посвящены теоретическим основам АСК-анализа и его практическим применениям в ряде предметных областей. На момент написания данной работы автором опубликовано более 40 монографий, 27 учебных пособий, в т.ч. 3 учебных пособия с грифами УМО и Министерства, получен 32 патента РФ на системы искусственного интеллекта, 352 публикации в изданиях, входящих в перечень ВАК РФ и приравненных им (по данным РИНЦ), 19 публикаций в изданиях, входящих в Ярор РИНЦ, 6 статей в журналах, входящих в WoS, 7 публикаций в журналах, входящих в Скопус⁸.

Три монографии включены в фонды библиотеки конгресса США⁹.

АСК-анализ и система "Эйдос" были успешно применены в 8 докторских диссертациях по экономическим, техническим, биологическим наукам и в 8 кандидатских диссертациях по экономическим, техническим, психологическими и медицинским наукам, еще четыре кандидатских диссертации с применением АСК-анализа находятся в стадии выхода на защиту.

⁶ http://lc.kubagro.ru/aidos/Sprab0802.pdf (см. с публикации № 48).

⁷ https://yandex.ru/search/?text=Aвтоматизированный%2Всистемно-когнитивный%2Ванализ%2В(АСК-анализ)&lr=35&clid=2327117-18&win=360

⁸ http://lc.kubagro.ru/aidos/Sprab0802.pdf

⁹ https://catalog.loc.gov/vwebv/search?searchArg=Lutsenko+E.V. (и кликнуть: "Search")

Автор является основателем междисциплинарной научной школы: «Автоматизированный системно-когнитивный анализ» 10. Научная школа: анализ" "Автоматизированный системно-когнитивный междисциплинарным научным направлением на пересечении по крайней мере трех научных специальностей (согласно недавно утвержденной новой номенклатуры научных специальностей ВАК $P\Phi^{11}$). Основные научные специальности, которым соответствует научная школа:

- 5.12.4. Когнитивное моделирование;
- 1.2.1. Искусственный интеллект и машинное обучение;
- 2.3.1. Системный анализ, управление и обработка информации.

школа: "Автоматизированный системно-когнитивный анализ" включает следующие междисциплинарные научные направления:

- Автоматизированный системно-когнитивный анализ числовых и текстовых табличных данных;
- Автоматизированный системно-когнитивный анализ *текстовых* данных;
- Спектральный и контурный автоматизированный системнокогнитивный анализ изображений;
- Сценарный автоматизированный системно-когнитивный анализ временных и динамических рядов.

Приводить здесь ссылки на все эти работы здесь вряд ли целесообразно. Отметим лишь, что у автора есть личный сайт [17] и страничка в РесечГейт [18], на которых можно получить более полную информацию о методе АСК-анализа и системе «Эйдос». Краткая информация об АСК-анализе и системе «Эйдос» есть в материале: http://lc.kubagro.ru/aidos/Presentation_Aidos-online.pdf.

Сопоставимость обработки числовых и текстовых представленных в разных единицах измерения, в АСК-анализе и системе «Эйдос» обеспечивается *путем* метризации номинальных шкал, т.е. повышения их степени формализации до уровня числовых шкал [21]. Сама метризация номинальных шкал достигается путем вычисления количества информации, содержащегося в градациях описательных шкал о получении тех или иных значений классификационных шкал. Для работы с лингвистическими переменными применяется лингвистический АСКанализ [22].

Суть метода АСК-анализа состоит в последовательном повышении степени формализации модели и преобразовании данных в информацию, а ее в знания и решении на основе этих знаний задач идентификации (распознавания, классификации и прогнозирования), поддержки принятия

https://www.famous-scientists.ru/school/1608
 https://www.garant.ru/products/ipo/prime/doc/400450248/

решений и исследования моделируемой предметной области (рисунки 1, 2) [19, 20].

О соотношении содержания понятий: «Данные», «Информация» и «Знания» **ДАННЫЕ**(ИНФОРМАЦИЯ, ЗАПИСАННАЯ НА КАКОМ-ЛИБО ЯЗЫКЕ ИЛИ В СИСТЕМЕ КОДИРОВАНИЯ НА НОСИТЕЛЕ ИЛИ НАХОДЯЩАЯСЯ В КАНАЛЕ СВЯЗИ И РАССМАТРИВАЕМАЯ БЕЗОТНОСИТЕЛЬНО К ЕЕ СМЫСЛОВОМУ СОДЕРЖАНИЮ) ИНФОРМАЦИЯ (ОСМЫСЛЕННЫЕ ДАННЫЕ, Т.Е. СОГЛАСНО КОНЦЕПЦИИ СМЫСЛА ШЕНКА-АБЕЛЬСОНА ДАННЫЕ, В КОТОРЫХ ВЫЯВЛЕНЫ ПРИЧИННО-СЛЕДСТВЕННЫЕ ЗАВИСИМОСТИ) Анализ данных выявление в них событий и причинноследственных связей **ЗНАНИЯ** Постановка целей, между событиями оценка знака и степени (ИНФОРМАЦИЯ, ПОЛЕЗНАЯ влияния информации ДЛЯ ДОСТИЖЕНИЯ ЦЕЛЕЙ, на их достижение Т.Е. УПРАВЛЕНИЯ)

Рисунок 1. О соотношении содержания понятий: «данные», «информация» и «знания» в АСК-анализе

<u>Данные</u> — это информация, записанная на каком-либо носителе или находящаяся в каналах связи и представленная на каком-то языке или в системе кодирования и рассматриваемая безотносительно к ее смысловому содержанию.

Исходные данные об объекте управления обычно представлены в форме баз данных, чаще всего временных рядов, т.е. данных, привязанных ко времени. В соответствии с методологией и технологией автоматизированного системно-когнитивного анализа (АСК-анализ),

развиваемой проф. Е.В.Луценко, для управления и принятия решений использовать непосредственно исходные данные не представляется возможным. Точнее сделать это можно, но результат управления при таком подходе оказывается мало чем отличающимся от случайного. Для реального же решения задачи управления необходимо предварительно преобразовать данные в информацию, а ее в знания о том, какие воздействия на корпорацию к каким ее изменениям обычно, как показывает опыт, приводят.

Информация есть осмысленные данные.

Смысл данных, в соответствии с концепцией смысла Шенка-Абельсона, состоит в том, что известны причинно-следственные зависимости между событиями, которые описываются этими данными. Таким образом, данные преобразуются в информацию в результате операции, которая называется «Анализ данных», которая состоит из двух этапов:

- 1. Выявление событий в данных (разработка классификационных и описательных шкал и градаций и преобразование с их использованием исходных данных в обучающую выборку, т.е. в базу событий эвентологическую базу).
- 2. Выявление причинно-следственных зависимостей между событиями.

В случае систем управления событиями в данных являются совпадения определенных значений входных факторов и выходных параметров объекта управления, т.е. по сути, случаи перехода объекта управления в определенные будущие состояния под действием определенных сочетаний значений управляющих факторов. Качественные значения входных факторов и выходных параметров естественно формализовать в форме лингвистических переменных. Если же входные факторы и выходные параметры являются числовыми, то их значения измеряются с некоторой погрешностью и фактически представляют собой интервальные числовые значения, которые также ΜΟΓΥΤ быть представлены или формализованы в форме лингвистических переменных «средние», «большие» значения «малые», экономических показателей).

Какие же математические меры могут быть использованы для количественного измерения силы и направления причинно-следственных зависимостей?

Наиболее очевидным ответом на этот вопрос, который обычно первым всем приходит на ум, является: «Корреляция». Однако, в статистике это хорошо известно, что это совершенно не так. Для преобразования исходных данных в информацию необходимо не только выявить события в этих данных, но и найти причинно-следственные связи

между этими событиями. В АСК-анализе предлагается 7 количественных мер причинно-следственных связей, основной из которых является семантическая мера целесообразности информации по А.Харкевичу.

Знания – это информация, полезная для достижения целей.

Значит для преобразования информации в знания необходимо:

- 1. Поставить цель (классифицировать будущие состояния моделируемого объекта на целевые и нежелательные).
- 2. Оценить полезность информации для достижения этой цели (знак и силу влияния).

Второй пункт, по сути, выполнен при преобразовании данных в информацию. Поэтому остается выполнить только первый пункт, т.к. классифицировать будущие состояния объекта управления как желательные (целевые) и нежелательные.

Знания могут быть представлены в различных формах, характеризующихся различной степенью формализации:

- вообще неформализованные знания, т.е. знания в своей собственной форме, ноу-хау (мышление без вербализации есть медитация);
 - знания, формализованные в естественном вербальном языке;
- знания, формализованные в виде различных методик, схем, алгоритмов, планов, таблиц и отношений между ними (базы данных);
- знания в форме технологий, организационных, производственных, социально-экономических и политических структур;
- знания, формализованные в виде математических моделей и методов представления знаний в автоматизированных интеллектуальных системах (логическая, фреймовая, сетевая, продукционная, нейросетевая, нечеткая и другие).

Таким образом, для решения сформулированной проблемы необходимо осознанно и целенаправленно **последовательно повышать степень формализации** исходных данных до уровня, который позволяет ввести исходные данные в интеллектуальную систему, а затем:

- преобразовать исходные данные в информацию;
- преобразовать информацию в знания;
- использовать знания для решения задач управления, принятия решений и исследования предметной области.

Последовательность обработки данных, информации и знаний в системе «Эйдос», повышение уровня системности данных, информации и знаний, повышение уровня системности моделей

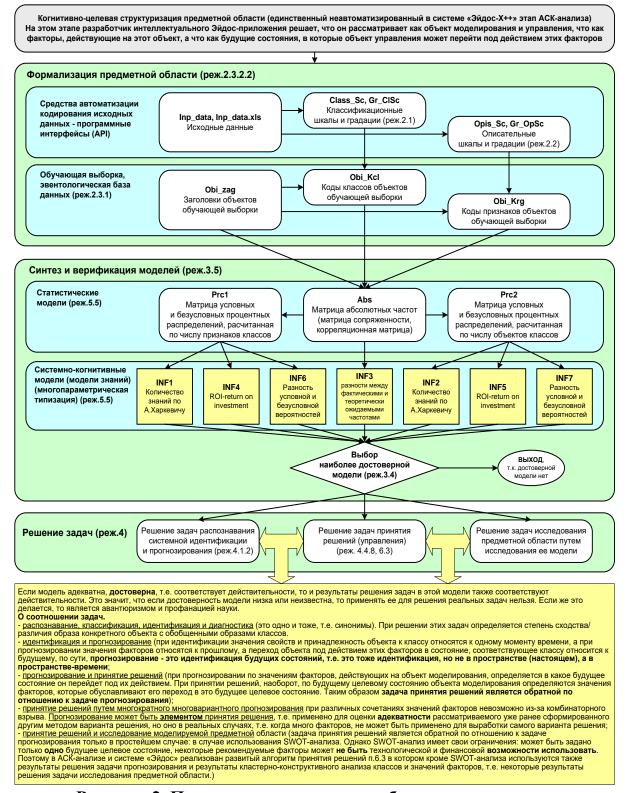


Рисунок 2. Последовательность преобразования данных в информацию, а ее в знания и решения задач в АСК-анализе и системе «Эйдос»

2.3.2. Суть математической модели АСК-анализа 2.3.2.1. Частные критерии

Математические модели, на основе которых рассчитываются статистические и системно-когнитивные модели автоматизированного системно-когнитивного анализа (АСК-анализ) и интеллектуальной системы «Эйдос», подробно описаны в ряде монографий и статей автора [19, 20]. Поэтому в данной работе мы рассмотрим эти вопросы очень кратко, акцентируя внимание лишь на математической взаимосвязи коэффициента возврата инвестиций (ROI) с мерой х-квадрат Карла Пирсона и с семантической мерой целесообразности информации Александра Харкевича.

Отметим, что модели системы «Эйдос» основаны на матрице абсолютных частот, отражающей число встреч градаций описательных шкал по градациям классификационных шкал (фактов). Но для решения всех задач используется не непосредственно сама эта матрица, а матрицы условных и безусловных процентных распределений и системно-когнитивные модели, которые рассчитываются на ее основе и отражают какое количество информации содержится в факте наблюдения определенной градации описательной шкалы о том, что объект моделирования перейдет в состояние, соответствующее определенной градации классификационной шкалы (классу).

Математическая модель АСК-анализа и системы «Эйдос» основана на системной нечеткой интервальной математике [20] и обеспечивает сопоставимую обработку больших объемов фрагментированных и зашумленных взаимозависимых (нелинейных) данных, представленных в различных типах шкал (дихотомических, номинальных, порядковых и числовых) и различных единицах измерения.

Пордок расчета частных критриев

Непосредственно на основе эмпирических данных рассчитывается матрица абсолютных частот (таблица 2).

На основе таблицы 2 рассчитываются матрицы условных и безусловных процентных распределений (таблица 3).

1-й способ: в качестве $N_{\Sigma j}$ используется суммарное количество признаков по классу;

2-й способ: в качестве $N_{\Sigma j}$ используется суммарное количество объектов обучающей выборки по классу.

Отметим, что в АСК-анализе и его программном инструментарии интеллектуальной системе «Эйдос» используется два способа расчета матриц условных и безусловных процентных распределений:

Таблица 1 – Матрица абсолютных частот (статистическая модель ABS)

Классы						подоль подо	
		1		i		W	Сумма
80	1	N_{11}	***	N_{1j}		N_{1W}	Сумма
bd							
Значения факторов	i	N_{i1}		N_{ij}		N_{iW}	$N_{i\Sigma} = \sum_{j=1}^W N_{ij}$
Har	•••						
m m	М	$N_{ m M1}$		N_{Mj}		$N_{\scriptscriptstyle MW}$	
	ммарное пичество о классу			$N_{\Sigma j} = \sum_{i=1}^M N_{ij}$			$N_{\Sigma\Sigma} = \sum_{i=1}^W \sum_{j=1}^M N_{ij}$
Суммарное количество объектов обучающей выборки по классу				$N_{\scriptscriptstyle \Sigma j}$			$N_{\Sigma\Sigma} = \sum_{j=1}^W N_{\Sigma j}$

Таблица 2 – Матрица условных и безусловных процентных распределений (статистические модели PRC1 и PRC2)

распределении (статистические модели РКСТ и РКС2)							
			Классы	Безусловная			
		1		j		W	вероятность признака
_	1	P_{11}		P_{1j}		P_{1W}	
00							
Значения факторов	i	P_{i1}		$P_{ij} = rac{N_{ij}}{N_{\Sigma j}}$		P_{iW}	$P_{i\Sigma} = rac{N_{i\Sigma}}{N_{\Sigma\Sigma}}$
3 1							
	М	$P_{ m M1}$		P_{Mj}		P_{MW}	
Безусло вероятн класс	ость			$P_{\Sigma j}$			

На практике часто встречается существенная несбалансированность данных, под которой понимается сильно отличающееся количество наблюдений объектов обучающейся выборки, относящихся к различным градациям одной классификационной или описательной шкалы. Поэтому решать задачу на основе непосредственно матрицы абсолютных частот

(таблица 2) было бы очень неразумно и переход от абсолютных частот к условным и безусловным относительным частотам (частостям) (таблица 3) является весьма обоснованным и логичным.

Этот переход полностью снимает *проблему несбалансированности данных*, т.к. в последующем анализе используется не матрица абсолютных частот (таблица 2), а матрицы условных и безусловных процентных распределений (таблица 3), а также матрицы системно-когнитивных моделей, рассчитываемые на основе матрица абсолютных частот и матрицы условных и безусловных процентных распределений. Этот подход снимает также проблему обеспечения сопоставимости обработки в одной модели исходных данных, представленных в различных видах шкал (дихотомических, номинальных, порядковых и числовых) и в разных единицах измерения [21]. В системе «Эйдос» этот подход применяется всегда при решении любых задач.

Затем на основе таблиц 2 и 3 с использованием частных критериев, знаний приведенных таблице 4, рассчитываются матрицы 7 системно-когнитивных моделей (таблица 5).

В таблице 4 приведены формулы:

- для сравнения фактических и теоретических абсолютных частот;
- для сравнения условных и безусловных относительных частот («вероятностей»).

И это *сравнение* в таблицах 2 и 3 осуществляется двумя возможными способами: путем *вычитания* и путем *деления*.

Количество частных критериев знаний и основанных на них системно-когнитивных моделей (таблица 4), применяемых в настоящее время в системе «Эйдос» равное 7 определяется тем, что они получаются путем всех возможных вариантов сравнения фактических и теоретических абсолютных частот, условных и безусловных относительных частот путем вычитания и путем деления, и при этом Nj рассматривается как суммарное количество или признаков, или объектов обучающей выборки в j-м классе, а нормировка к нулю (для аддитивных интегральных критериев), если нет связи между наличием признака и принадлежностью объекта к классу, осуществляется либо логарифмированием, либо вычитанием единицы (таблица 6).

Таблица 3– Различные аналитические формы частных критериев знаний, применяемые в АСК-анализе и системе «Эйдос»

применяемые в АСК-анализе и	ONOTONO WO	<u>пдоои</u>
	Выражение	для частного критерия
Наименование модели знаний и частный критерий	Через относительны е частоты	Через абсолютные частоты
ABS , матрица абсолютных частот, <i>Nij</i> - фактическое число	$N_{ij} - \phi$ акті	ическая частота;
встреч <i>i-го</i> признака у объектов <i>j-го</i> класса; \overline{N}_{ij} - теоретическое число встреч <i>i-го</i> признака у объектов <i>j-го</i> класса; Ni – суммарное количество признаков в i - \check{u} строке; Nj	$N_i = \sum_{j=1}^W N_{ij};$	$N_{j} = \sum_{i=1}^{M} N_{ij}; N = \sum_{i=1}^{W} \sum_{j=1}^{M}$
– суммарное количество признаков или объектов обучающей выборки в <i>ј-м</i> классе; <i>N</i> – суммарное количество признаков по всей выборке (таблица 7)	$\overline{N}_{ij} = \frac{N_i N_j}{N}$	– теоретическая част
PRC1, матрица условных <i>Pij</i> и безусловных <i>Pi</i> процентных распределений, в качестве <i>Nj</i> используется суммарное количество признаков по классу PRC2, матрица условных <i>Pij</i> и безусловных <i>Pi</i> процентных распределений, в качестве <i>Nj</i> используется суммарное количество объектов обучающей выборки по классу		$P_{ij} = \frac{N_{ij}}{N_j}; P_i = \frac{N_i}{N}$
INF1, частный критерий: количество знаний по А.Харкевичу, 1-й вариант расчета вероятностей: Nj − суммарное количество признаков по j-му классу. Вероятность того, что если у объекта j-го класса обнаружен признак, то это i-й признак INF2, частный критерий: количество знаний по А.Харкевичу, 2-й вариант расчета вероятностей: Nj − суммарное количество объектов по j-му классу. Вероятность того, что если предъявлен объект j-го класса, то у него будет обнаружен i-й признак.	$I_{ij} = \Psi \times Log_2$	$I_{ij} = \Psi \times Log_2 \frac{N_{ij}}{\overline{N}_{ij}} = \Psi$
INF3, частный критерий: Хи-квадрат: разности между фактическими и теоретически ожидаемыми абсолютными частотами		$I_{ij} = N_{ij} - \overline{N}_{ij} = N_{ij} - \frac{N}{N}$
INF4, частный критерий: ROI - Return On Investment, 1-й вариант расчета вероятностей: Nj – суммарное количество признаков по j-му классу INF5, частный критерий: ROI - Return On Investment, 2-й вариант расчета вероятностей: Nj – суммарное количество объектов по j-му классу	$I_{ij} = \frac{P_{ij}}{P_i} - 1 = \frac{P_{ij}}{P_i}$	$I_{ij} = \frac{N_{ij}}{\overline{N}_{ij}} - 1 = \frac{N_{ij}N_{ij}}{N_{i}N_{ij}}$
INF6, частный критерий: разность условной и безусловной вероятностей, 1-й вариант расчета вероятностей: Nj — суммарное количество признаков по j-мy классу INF7, частный критерий: разность условной и безусловной вероятностей, 2-й вариант расчета вероятностей: Nj — суммарное количество объектов по j-мy классу	$I_{ij} = P_{ij} - P_i$	$I_{ij} = \frac{N_{ij}}{N_j} - \frac{N_i}{N} = \frac{N_{ij}N - N_i}{N_jN_i}$

Обозначения к таблице:

і – значение прошлого параметра;

ј - значение будущего параметра;

Nij – количество встреч j-го значения будущего параметра при i-м значении прошлого параметра;

М – суммарное число значений всех прошлых параметров;

W - суммарное число значений всех будущих параметров.

Ni – количество встреч i-м значения прошлого параметра по всей выборке:

Nj – количество встреч j-го значения будущего параметра по всей выборке;

N – количество встреч j-го значения будущего параметра при i-м значении прошлого параметра по всей выборке.

lij – частный критерий знаний: количество знаний в факте наблюдения i-го значения прошлого параметра о том, что объект перейдет в состояние, соответствующее j-му значению будущего параметра;

 Ψ – нормировочный коэффициент (Е.В.Луценко, 2002), преобразующий количество информации в формуле А.Харкевича в биты и обеспечивающий для нее соблюдение принципа соответствия с формулой Р.Хартли;

Pi – безусловная относительная частота встречи i-го значения прошлого параметра в обучающей выборке:

Pij – условная относительная частота встречи i-го значения прошлого параметра при j-м значении будущего параметра.

Таблица 4 – Матрица системно-когнитивной модели

	таолица 4 — матрица системно-когнитивной модели									
				Классь	ol		Значимость			
		1		j		W	фактора			
08	1	I_{11}		I_{1j}		I_{1W}	$\sigma_{1\Sigma} = \sqrt[2]{\frac{1}{W-1} \sum_{j=1}^{W} (I_{1j} - \bar{I}_1)^2}$			
<u></u> 6										
Значения факторов	i	I_{i1}		I_{ij}		I_{iW}	$\sigma_{i\Sigma} = \sqrt[2]{\frac{1}{W-1} \sum_{j=1}^{W} \left(I_{ij} - \bar{I}_{i}\right)^{2}}$			
<u>ā</u>										
Ė	М	$I_{ m M1}$		I_{Mj}		I_{MW}	$\sigma_{M\Sigma} = \sqrt[2]{\frac{1}{W-1} \sum_{j=1}^{W} \left(I_{Mj} - \bar{I}_{M} \right)^{2}}$			
Степ редукі кла		$\sigma_{\scriptscriptstyle{\Sigma1}}$		$\sigma_{\scriptscriptstyle{\Sigma}\! j}$		$\sigma_{\scriptscriptstyle \Sigma W}$	$H = \sqrt[2]{\frac{1}{(W \cdot M - 1)} \sum_{j=1}^{W} \sum_{i=1}^{M} (I_{ij} - \bar{I})^{2}}$			

Таблица 5– Конфигуратор системно-когнитивных моделей АСК-анализа и интеллектуальной системы «Эйдос»

	Способ сравнения	Нормировка не требуется	Нормировка к 0 путем взятия логарифма	Нормировка к 0 путем вычитания 1
Сравнение фактических и теоретических	Путем деления		INF1, INF2, Александра Харкевича	INF4, INF5, Коэффициент возврата инвестиций ROI
абсолютных частот	Путем вычитания	INF3, χ-квадрат Карла Пирсона		
Сравнение условных и безусловных	Путем деления		INF1, INF2, Александра Харкевича	INF4, INF5, Коэффициент возврата инвестиций ROI
относительных частот	Путем вычитания	INF6, INF7		

Обратим особое внимание на то, что сравнение фактических и теоретических абсолютных частот путем деления приводит (что применения нормировках К НУЛЮ нужно ДЛЯ аддитивных интегральных критериев) путем взятия логарифма и путем вычитания 1 к тем же самым моделям, что и сравнение условных и безусловных относительных частот путем деления с теми же самыми способами нормировки. Таким образом, если на основе матрицы абсолютных частот рассчитать матрицы условных и безусловных процентных распределений, а затем сравнить фактические абсолютные частоты с теоретическими путем вычитания и деления, а также сравнить условные и безусловные относительные частоты также путем вычитания и деления и провести нормировку к 0 путем взятия логарифма и путем вычитания 1, то получается 3 статистических модели: матрица абсолютных частот и две матрицы относительных частот, т.е. условных и безусловных процентных распределений, а также всего 7 системно-когнитивных моделей. Других же системно-когнитивных моделей, рассчитываемых на основе приведенных нет. моделей просто Это И есть конфигуратор статистических и когнитивных моделей в смысле В.А.Лефевра. Под конфигуратором В.А.Лефевр понимал минимальный полный набор понятийных шкал или конструктов, т.е. понятий, достаточный для адекватного описания предметной области $[6]^{12}$. Необходимо отметить, что все эти модели рассчитываются в интеллектуальной системе «Эйдос».

Когда мы сравниваем фактические и теоретические абсолютные частоты путем вычитания у нас получается частный критерий знаний: «хиквадрат» (СК-модель INF3), когда же мы сравниваем их путем деления, то у нас получается частный критерий: «количество информации по А.Харкевичу» (СК-модели INF1, INF2) или «коэффициент возврата инвестиций ROI» - Return On Investment (СК-модели INF4, INF5) в зависимости от способа нормировки.

Когда же мы сравниваем условные и безусловные относительные частоты путем вычитания у нас получается частный критерий знаний: «коэффициент взаимосвязи» (СК-модели INF6, INF7), когда же мы сравниваем их путем деления, то у нас получается частный критерий: «количество информации по A.Харкевичу» (СК-модели INF1, INF2).

Таким образом, мы видим, что все частные критерии знаний тесно взаимосвязаны друг с другом. Особенно интересна связь знаменитого критерия хи-квадрат К.Пирсона с замечательной мерой количества информации А.Харкевича и с известным в экономике коэффициентом ROI.

Вероятность рассматривается как предел, к которому стремится относительная частота (отношение количества благоприятных исходов к числу испытаний) при неограниченном увеличении количества испытаний. Ясно, что вероятность — это математическая абстракция, которая никогда не встречается на практике (также как и другие математические и физические абстракции, типа математической точки, материальной точки, бесконечно малой и т.п.). На практике встречается только относительная частота. Но она может быть весьма близкой к вероятности. Например, при 480 наблюдений различие между относительной частой и вероятностью (погрешность) составляет около 5%, при 1250 наблюдениях — около 2.5%, при 10000 наблюдениях — 1%.

Суть этих методов в том, что вычисляется количество информации в значении фактора о том, что объект моделирования перейдет под его

12 См. 1.2.1.2.1.1. Определение понятия конфигуратора, http://lc.kubagro.ru/aidos/aidos06_lec/index.htm

19

действием в определенное состояние, соответствующее классу. Это позволяет сопоставимо и корректно обрабатывать разнородную информацию о наблюдениях объекта моделирования, представленную в различных типах измерительных шкал и различных единицах измерения [21].

На основе системно-когнитивных моделей, представленных в таблице 5 (отличаются частыми критериями, приведенными в таблице 4), решаются задачи идентификации (классификации, распознавания, диагностики, прогнозирования), поддержки принятия решений (обратная задача прогнозирования), а также задача исследования моделируемой предметной области путем исследования ее системно-когнитивной модели.

Отметим, что как значимость значения фактора, степень детерминированности класса и ценность или качество модели в АСК-анализе рассматривается вариабельность значений частных критериев этого значения фактора, класса или модели в целом (таблица 6).

Численно эта вариабельность может измеряться разными способами, например средним отклонением модулей частных критериев от среднего, дисперсией или среднеквадратичным отклонением или его квадратом. В системе «Эйдос» принят последний вариант, т.к. эта величина совпадает с мощностью сигнала, в частности мощностью информации, а в АСК-анализе все модели рассматриваются как источник информации об объекте моделирования. Поэтому есть все основания уточнить традиционную терминологию АСК-анализа (таблица 7):

Таблица 6 – Уточнение терминологии АСК-анализа

		Новый	
Nº	Традиционные термины (синонимы)	термин	Формула
1	1. Значимость значения фактора (признака). 2. Дифференцирующая мощность значения фактора (признака). 3. Ценность значения фактора (признака) для решения задачи идентификации и других задач	Корень из информационной мощности значения фактора	$\sigma_{i\Sigma} = \sqrt[2]{\frac{1}{W-1} \sum_{j=1}^{W} \left(I_{ij} - \bar{I}_i\right)^2}$
2	Степень детерминированности класса. Степень обусловленности класса.	Корень из информационной мощности класса	$\sigma_{\Sigma j} = \sqrt[2]{\frac{1}{M-1} \sum_{i=1}^{M} \left(I_{ij} - \bar{I}_{j}\right)^{2}}$
3	 Качество модели. Ценность модели. Степень сформированности модели. Количественная мера степени выраженности закономерностей в моделируемой предметной области 	Корень из информационной мощности модели	$H = \sqrt[2]{\frac{1}{(W \cdot M - 1)} \sum_{j=1}^{W} \sum_{i=1}^{M} (I_{ij} - \bar{I})^{2}}$

Итак, в данном разделе раскрывается простая Математическая взаимосвязь меры χ-квадрат Карла Пирсона с коэффициентом возврата инвестиций (ROI) и с семантической мерой целесообразности информации

Александра Харкевича. Эта взаимосвязь обнаруживается, если на основе матрицы абсолютных частот рассчитать матрицы условных и безусловных процентных распределений, а затем сравнить фактические абсолютные частоты с теоретическими путем вычитания и деления, а также сравнить условные и безусловные относительные частоты также путем вычитания и деления и выполнить нормировку к нулю путем взятия логарифма или вычитания 1. При этом получается 3 статистических модели: матрица абсолютных частот и две матрицы относительных частот, т.е. условных и безусловных процентных распределений, а также всего 7 системнокогнитивных моделей. Именно 7, а не большее количество системномоделей в итоге получается потому, получающиеся в результате сравнения фактических и теоретических абсолютных частот путем деления и нормировки к нулю путем взятия логарифма или вычитания 1 тождественно совпадают с моделями, получающимися путем сравнения условных и безусловных относительных частот путем деления и нормировки к нулю путем взятия логарифма или вычитания 1. Это и есть конфигуратор статистических и когнитивных моделей в смысле В.А.Лефевра, содержащий минимальное количество моделей, позволяющих полно описать моделируемую предметную область.

Показательно, что модель меры х-квадрат Карла Пирсона из статистики оказалась математически тесно связанной с коэффициентом возврата инвестиций (ROI), применяемой в экономике в теории управления портфелем инвестиций и с мерой информации Александра Харкевича из семантической теории информации и теории управления знаниями. Все эти модели рассчитываются в интеллектуальной системе «Эйдос».

2.3.2.2. Интегральные критерии

2.3.2.2.1. Для чего нужны интегральные критерии: для системной идентификации и прогнозирования

Из статистических и системно-когнитивных моделей, которые создаются и верифицируются в режиме 3.5 системы «Эйдос» и отображаются в режиме 5.5, мы узнаем, какое количество информации о принадлежности или непринадлежности объекта к классу мы получаем из факта наблюдения у этого объекта одного некоторого признака (значения свойства или значения фактора, т.е. градации описательной шкалы). Но у объекта может быть не один, а много признаков. Как тогда определить одним числом степень сходства/различия этого объекта с классами, обобщающим информацию об этом, содержащуюся во всей системе признаков объекта? Интегральные критерии и представляют собой ответ на этот вопрос.

При решении задачи идентификации (задачи-6 АСК-анализа) каждый объект распознаваемой выборки сравнивается по всем своим признакам с каждым из обобщенных образов классов. *Смысл* решения задачи идентификации заключается в том, что при определении принадлежности конкретного объекта к обобщенному образу классу об этом конкретном объекте *по аналогии* становится известно все, что известно об объектах этого класса, по крайней мере, самое существенное о них, т.е. чем они отличаются от объектов других классов.

Задачи идентификации и прогнозирования взаимосвязаны и мало чем отличаются друг от друга. Главное различие между ними в том, что при идентификации значения свойств и принадлежность объекта к классу относятся к одному моменту времени, а при прогнозировании значения факторов относятся к прошлому, а переход объекта под действием этих факторов в состояние, соответствующее классу относится к будущему.

Задача решается в модели, заданной в качестве текущей, т.к. является весьма трудоемкой в вычислительном отношении. Правда с разработкой и реализацией в системе «Эйдос» высокоэффективных алгоритмов распознавания и использованием графического процессора (GPU) для расчетов эта проблема практически снялась.

Сравнение конкретного образа объекта с обобщенными образами классов осуществляется путем применения *неметрических интегральных критериев*, которых в настоящее время в системе «Эйдос» используется два. Эти интегральные критерии интересны тем, что корректны в неортонормированных пространствах, которые всегда и встречаются на практике, и являются фильтрами подавления шума.

2.3.2.2.2. Интегральный критерий «Сумма знаний»

Данный интегральный критерий представляет собой суммарное количество знаний, содержащееся в системе факторов различной природы, характеризующих сам объект управления, управляющие факторы и окружающую среду, о переходе объекта в будущие целевые или нежелательные состояния.

Интегральный критерий представляет собой аддитивную функцию от <u>частных критериев знаний</u>, представленных в help режима 5.5:

$$I_j = (\vec{I}_{ij}, \vec{L}_i).$$

В выражении круглыми скобками обозначено скалярное произведение. В координатной форме это выражение имеет вид:

$$I_{j} = \sum_{i=1}^{M} I_{ij} L_{i},$$

где: M – количество градаций описательных шкал (признаков);

 $ec{I}_{ij} = \{I_{ij}\}_{-\ {
m Beктор}}$ состояния ј—го класса;

 $\vec{L}_i = \{L_i\}$ — вектор состояния распознаваемого объекта, включающий все виды факторов, характеризующих сам объект, управляющие воздействия и окружающую среду (массив—локатор), т.е.:

$$\vec{L}_i = \begin{cases} 1, \textit{если } i - \breve{u} \ \phi \textit{актор действует}; \\ n, \textit{где} : n > 0, \textit{если } i - \breve{u} \ \phi \textit{актор действует с истин н остью } n; \\ 0, \textit{если } i - \breve{u} \ \phi \textit{актор н е действует}. \end{cases}$$

В текущей версии системы «Эйдос-X++» значения координат вектора состояния распознаваемого объекта принимались равными либо 0, если признака нет, или п, если он присутствует у объекта с интенсивностью п, т.е. представлен п раз (например, буква «о» в слове «молоко» представлена 3 раза, а буква «м» - один раз).

2.3.2.2.3. Интегральный критерий «Семантический резонанс знаний»

Этот интегральный критерий представляет собой нормированное суммарное количество знаний, содержащееся в системе факторов различной природы, характеризующих сам объект управления, управляющие факторы и окружающую среду, о переходе объекта в будущие целевые или нежелательные состояния.

Интегральный критерий представляет собой аддитивную функцию от <u>частных критериев знаний</u>, представленных в help режима 3.3 и имеет вид:

$$I_{j} = \frac{1}{\sigma_{i}\sigma_{l}M} \sum_{i=1}^{M} \left(I_{ij} - \bar{I}_{j}\right) \left(L_{i} - \bar{L}\right),$$

где:

M — количество градаций описательных шкал (признаков); \bar{I}_j — средняя информативность по вектору класса; \bar{L} — среднее по вектору объекта;

 σ_{j} — среднеквадратичное отклонение частных критериев знаний вектора класса; σ_{l} — среднеквадратичное отклонение по вектору распознаваемого объекта.

 $\vec{I}_{ij} = \{I_{ij}\}$ — вектор состояния j—го класса; $\vec{L}_i = \{L_i\}$ — вектор состояния распознаваемого объекта (состояния или явления), включающий все виды факторов, характеризующих сам объект, управляющие воздействия и окружающую среду (массив—локатор), т.е.:

$$\vec{L}_{i} = \begin{cases} 1, \textit{если } i - \breve{u} \; \phi \textit{актор действует}; \\ n, \textit{где} : n > 0, \textit{если } i - \breve{u} \; \phi \textit{актор действует c истин н оство n}; \\ 0, \textit{если } i - \breve{u} \; \phi \textit{актор н е действует}. \end{cases}$$

В текущей версии системы «Эйдос-X++» значения координат вектора состояния распознаваемого объекта принимались равными либо 0, если признака нет, или n, если он присутствует у объекта с интенсивностью n, т.е. представлен n раз (например, буква «о» в слове «молоко» представлена 3 раза, а буква «м» - один раз).

«молоко» представлена 3 раза, а буква «м» - один раз). выражение Приведенное ДЛЯ критерия «Семантический резонанс знаний» получается непосредственно из выражения для критерия «Сумма знаний» после замены координат перемножаемых векторов их стандартизированными значениями: $I_{ij} \to \frac{I_{ij} - \bar{I}_{j}}{\sigma_{i}}, \ L_{i} \to \frac{L_{i} - \bar{L}}{\sigma_{i}}$. Поэтому по своей сути он также является скалярным произведением двух стандартизированных (единичных) векторов класса и объекта. Существуют и много других способов нормирования, например, путем применяя сплайнов, в частности линейной интерполяции: $I_{ij} o rac{I_{ij} - I_{j}^{\min}}{I_{i}^{\max} - I_{i}^{\min}}, \;\; L_{i} o rac{L_{i} - L^{\min}}{L^{\max} - L^{\min}}, \;$ Это позволяет предложить неограниченное количество других видов интегральных критериев. Но результаты их применения едва ли они будут существенно отличаться существующих, поэтому они не реализованы в системе «Эйдос».

2.3.2.2.4. Некоторые полезные математические свойства интегральных критериев

Данные интегральные критерии обладают очень интересными и полезными *математическими свойствами*, которые обеспечивают ему важные достоинства и преимущества по сравнению с другими интегральными критериями:

<u>Во-первых</u>, интегральный критерий имеет **неметрическую** природу, т.е. он являются мерой сходства векторов класса и объекта, но не расстоянием между ними, а косинусом угла между ними, т.е. это межвекторное или информационное расстояние. Поэтому его применение является корректным в **неортонормированных** пространствах, которые, как правило, и встречаются на практике и в которых применение Евклидова расстояния (теоремы Пифагора) является некорректным.

<u>Во-вторых</u>, данный интегральный критерий являются фильтром, подавляющим белый шум, который всегда присутствует в эмпирических исходных данных и в моделях, созданных на их основе. Это свойство подавлять белый шум проявляется у данного критерия тем ярче, чем больше в модели градаций описательных шкал.

<u>В-третьих</u>, интегральный критерий сходства представляет собой количественную меру сходства/различия конкретного объекта с обобщенным образом класса и имеет тот же смысл, что и функция принадлежности элемента множеству в нечеткой логике Лотфи Заде. Однако в нечеткой логике эта функция задается исследователем априорно путем выбора из нескольких возможных вариантов, а в АСК-анализе и его программном инструментарии — интеллектуальной системе «Эйдос» она рассчитывается в соответствии с хорошо обоснованной математической моделью непосредственно на основе эмпирических данных.

<u>В-чемвертых</u>, кроме того значение интегрального критерия сходства представляет собой адекватную самооценку **степени уверенности** системы в положительном или отрицательном решении о принадлежности/непринадлежности объекта к классу или **риска ошибки** при таком решении.

B-пятых, по сути, при распознавании происходит расчет коэффициентов I_j разложения функции объекта L_i в ряд по функциям классов I_{ij} , т.е. определяется **вес** каждого обобщенного образа класса в образе объекта, что подробнее описано в монографии [23].

2.3.2.3. Модели АСК-анализа как матричные функции отображения пространства факторов на пространство классов (будущих состояний объекта моделирования)

2.3.2.3.1. Конструкты и когнитивные пространства факторов и классов

Конструктом называется понятие, имеющее противоположные по смыслу полюса и спектр промежуточных смысловых значений, расположенных в некоторой шкале: порядковой или числовой [24].

Например, понятие «температура» имеет полюса «горячее» и «холодное» с промежуточными значениями «теплое» и «прохладное» и шкалой Цельсия для количественного измерения температуры. Другие примеры конструктов: «вес: легкий-тяжелый», «размер: меленький-большой», «возраст: молодой-старый», «цвет: красный-фиолетовый», «нота: до-си» и т.д.

Часто бывает, что конструкт четко выражен своими полюсами и спектром, которые известны, но не имеет своего названия в языке. Например: «Дух-материя», «добро-зло», «истина-ложь».

Одним из наиболее фундаментальных свойств человеческого интеллекта является свойство мыслить в системе $\emph{бинарных}$ конструктами, т.е. конструктов, т.е. конструктов с $\emph{двумя}$ полюсами.

Как правило, люди мыслят именно самими полюсами конструктов, а промежуточные смысловые оттенки используют реже, как бы видят мир в черно-белой палитре или градациях серого.

Для классической логики, основанной на логике Аристотеля, характерно наличие лишь двух значений истинности: «Истина и ложь». В нечеткой логике может быть много дискретных значений истинности, когда степень истинности обозначается рациональным или даже целым числом, а может быть и континуум значений истинности, когда степень истинности обозначится действительным числом.

Конструкты представляют собой оси координат когнитивного пространства, определяющего мировоззрение человека или модель определенной предметной области.

Факторам, т.е. описательным шкалам, соответствуют оси-конструкты некоторого пространства, которое мы назовем *пространством факторов*.

Будущие состояния объекта моделирования будем описывать с помощью классификационных шкал, значения которых являются классами и соответствуют конкретным будущим состояниям объекта моделирования, т.е. классам. Классификационным шкалам соответствуют оси-конструкты некоторого пространства, которое мы назовем *пространством классов*.

У разных людей в когнитивном пространстве различный *набор* конструктов, различное их *количество* и у конструктов различный *диапазон*, а также разная *точность* позиционирования градаций между полюсами.

У людей широко образованных, больше набор конструктов, чем у людей необразованных. Даже если одни и те же конструкты есть и у образованных, и у необразованных людей, то у образованных они имеют больший диапазон, чем у необразованных и более высокую точность измерения положения градаций между полюсами.

Например, конструкт «температура» в быту имеет полюса: холодное-горячее, с диапазоном между полюсами около 300 градусов и порядковой шкалой, а у физиков он имеет диапазон от абсолютного нуля Кельвина до десятков триллионов градусов (получено на БАК) с числовой шкалой.

Таким образом, у разных людей различная размерность и объем когнитивного пространства, а также различная его дискретность: у людей широко образованных, больше размерность и объем когнитивного пространства, а также точность позиционирования в нем, чем у людей необразованных

Кратко рассмотрим варианты отношения когнитивных пространств разных людей и *принцип Уильяма Росса Эшби*. Для этого можно использовать иллюстрацию отношений между множествами, приведенную на рисунке 2.

Объединение когнитивных пространств двух людей порождает новое когнитивное пространство, включающее все конструкты и первого, и второго пространства («одна голова хорошо, а две лучше»).

Пересечение когнитивных пространств двух людей порождает новое когнитивное пространство, включающее только конструкты общие для их когнитивных пространств. Это пересечение содержит общую понятийную базу, основу для взаимопонимания.

Разность когнитивных пространств двух людей порождает новое когнитивное пространство, содержащее только те конструкты, которые есть у первого из них, но отсутствуют у второго. Это то, что отражает различие в мировоззрениях этих двух людей.

Симметрическая разность когнитивных пространств двух людей порождает новое когнитивное пространство, содержащее только те конструкты, которые есть только у одного из них, но отсутствуют те, которые есть у обоих. Это то, что наиболее полно и ярко отражает различия в мировоззрениях этих двух людей.



Рисунок 1.

Отметим в этой связи, что в процессе эволюции сознания также изменяется объем и размерность когнитивного пространства, в частности изменяется диапазон конструкта: «объективное-субъективноенесуществующее» и, соответственно, содержание областей объективное, субъективное и несуществующее (рисунок 3) [13].



Рисунок 2. Периодическая критериальная классификация форм сознания (Луценко Е.В., 1978) [13]

В общем если сказать упрощенно, это приводит к тому, что при повышении уровня сознания объем осознаваемого постоянно увеличивается. Это означает, что человек в более высокой форме сознания будет иметь более адекватные и более полные модели реальности (себя и окружающего), чем человек в более низкой форме сознания.

Но этот процесс увеличения области осознаваемого происходит не просто путем объединения областей осознаваемого при предыдущих формах сознания, а представляет собой сложный диалектический процесс. Дело в том, что при разных формах сознания сами закономерности объективного, субъективного изменения осознания как несуществующего окружающей среды И самого себя отличается. Подробнее эти закономерности описаны в ряде работ [25].

В полном соответствии с принципом Уильяма Росса Эшби если общаются два человека с разными размерностями и объемами когнитивных пространств, то первый более развитый человек с большими размерностями и объемом когнитивного пространства будет адекватно осознавать второго человека с меньшими размерностями и объемом когнитивного пространства. А второй наоборот, будет осознавать первого упрощенно, ограниченно, ущербно, в проекции в свое когнитивное

пространство меньшего числа измерений, т.е. по сути, будет судить о нем по себе.

В этой связи уместно процитировать высказывание одного известного специалиста по когнитивным пространствам Козьмы Пруткова: «Многие вещи нам непонятны не потому, что наши понятия слабы, а потому, что сии вещи вообще не входят в круг наших понятий»).

Приходит на ум также буддийская притча об ученике, который с юности безуспешно искал Учителя и уже в зрелом возрасте наконец нашел его и им оказался его сосед, с которым он был знаком с детства, видел как он возится по хозяйству в своем дворе, но не видел в нем Учителя.

2.3.2.3.2. Когнитивные функции как nd-отображение пространства факторов на пространство классов

Будем считать, что нам удалось смоделировать объект моделирования, если мы определили силу и направление влияния каждого значения фактора на поведение объект моделирования, т.е. на его переходы в будущие состояния, соответствующие классам.

Представим себе, что факторам, т.е. описательным шкалам, соответствуют оси-конструкты некоторого пространства, которое мы назовем *пространством факторов*.

Будущие состояния объекта моделирования будем описывать с помощью классификационных шкал, значения которых являются классами и соответствуют конкретным будущим состояниям объекта моделирования, т.е. классам. Классификационным шкалам соответствуют оси-конструкты некоторого пространства, которое мы назовем пространством классов.

Таким образом, для создания модели детерминации поведения объекта моделирования нам необходимо найти закон отображения пространства значений факторов, действующих на объект моделирования, на пространство классов, т.е. пространство будущих состояний объекта моделирования.

Подобные отображения называются многомерными *функциями*, а само отображение в случае евклидова пространства конформным отображением или преобразованием (рисунок 3).

Однако для системно-когнитивного моделирования хорошо разработанный математический аппарат непрерывных функций и конформных отображений не является вполне адекватными средством.

Причины этого просты и очевидны.

Этот аппарат разработан для метрических ортонормированных пространств, т.е. таких пространств, оси координат которых являются числовыми шкалами и взаимно перпендикулярны друг другу, причем размерность этих пространств должна быть одинаковой.

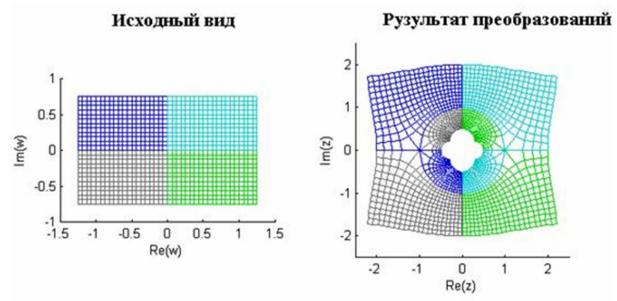


Рисунок 3. Конформное преобразование (сетка линий)¹³

Если говорить о пространстве факторов и пространстве состояний объекта моделирования, то это означает, что все они должны измеряться в количественных единицах измерения и некоторых должны быть независимы друг от друга, т.е. объект моделирования должен быть линейным [26] а количество описательных классификационных шкал должно быть одинаковым.

В нашем же случае пространство факторов и пространство классов, т.е. состояний объекта моделирования, в общем случае являются неметрическими неортонормированными пространствами, т.е. могут иметь оси, являющиеся не только числовыми шкалами, но и порядковыми, и номинальными [21], оси могут быть И ЭТИ перпендикулярными друг к другу, а значит взаимозависимыми, и количество этих осей в пространстве факторов и пространстве состояний может быть разным.

Математическое отображений моделирование неметрических неортонормированных пространств разной размерности друг на друга является математической проблемой. В теории линейных и конформных преобразований предполагается, что все оси пространств являются взаимно-перпендикулярными, а пространства одинаковую размерность.

В нашем же случае:

- факторы могут измеряться в разных типах описательных шкал и единицах измерения;
- будущие состояния объекта моделирования также могут измеряться в разных типах классификационных шкал и в разных единицах измерения;

¹³ Источник: http://matlab.exponenta.ru/imageprocess/book2/images-38/image004.jpg

- количество описательных и классификационных шкал может быть не равным друг другу;
- описательные шкалы могут быть взаимозависимы, и классификационные шкалы также могут быть взаимозависимы.

Поэтому в данной работе предлагается математическое преобразование, которое мы назовем обобщенным конформным отображением.

Для этого необходимо преобразовать номинальные и порядковые шкал в один тип шкал: числовой, и всех шкал в одну единицу измерения. Это преобразование измерительных шкал называется «метризация» [21].

В качестве этой единицы измерения нами выбрана единица измерения количества информации. Иначе говоря, в математической модели АСКанализа мы рассчитываем, какое количество информации содержится в различных значениях факторов о том, что объект моделирования перейдет в те или иные будущие состояния [21]. Причем делается это непосредственно на основе эмпирических данных, которые могут быть большой размерности, фрагментированы, зашумлены и взаимозависимы.

В пространстве факторов объект моделирования может быть представлен некоторой областью, соответствующей действующим на него значениям факторов.

В модели отражено количество информации о переходе объекта моделирования в каждое из будущих состояний содержится в каждом значении каждого фактора.

С помощью рассмотренных выше интегральных критериев мы можем количественно оценить совместное влияние всех значений факторов на переход объекта моделирования в будущие состояния, соответствующие классам.

Для каждой области в пространстве классов мы можем отобразить количество информации, которое содержится в системе значений факторов о переходе объекта моделирования в состояния, соответствующие этой области. На взгляд автора это наиболее удобно делать в виде линий разного цвета и толщины. Цвет линии будет означать знак, а толщина – модуль силы влияния.

Таким образом, если соединить каждую область в пространстве факторов с каждой из областей в пространстве классов линиями разного цвета и толщины, то это и будет *отображение пространства факторов на пространство классов*, отражающее в наглядном виде созданную модель поведения объекта моделирования под действием различных сочетаний значений факторов.

В интеллектуальной системе «Эйдос», являющейся в настоящее время программным инструментарием автоматизированного системно-

когнитивного анализа, реализован режим визуализации 2d-когнитивных функций (режим 4.5) [10].

2.3.3. Интеллектуальная система «Эйдос» – программный инструментарий АСК-анализа

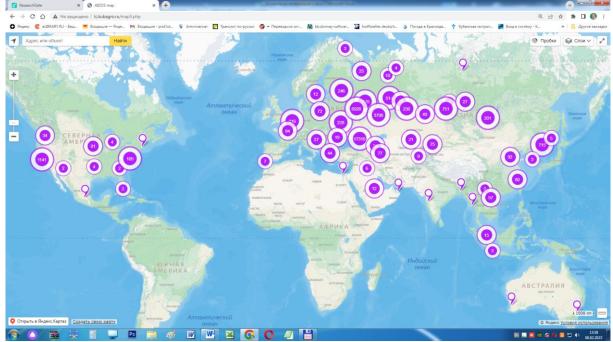
Существует много систем искусственного интеллекта. Универсальная когнитивная аналитическая система «Эйдос» отличается от них следующими параметрами:

- является универсальной и может быть применена во многих предметных областях, т.к. разработана в универсальной постановке, не зависящей от предметной области (http://lc.kubagro.ru/aidos/index.htm). Система «Эйдос» является автоматизированной системой, т.е. предполагает непосредственное участие человека в реальном времени при решении задач идентификации, прогнозирования, приятия решений и исследования предметной области (автоматические системы работают без такого участия человека);
- находится в полном открытом бесплатном доступе (http://lc.kubagro.ru/aidos/_Aidos-X.htm), причем с актуальными исходными текстами (http://lc.kubagro.ru/_AidosALL.txt): открытая лицензия: CC BY-SA 4.0 (https://creativecommons.org/licenses/by-sa/4.0/), и это означает, что ей могут пользоваться все, кто пожелает, без какого-либо дополнительного разрешения со стороны первичного правообладателя автора системы «Эйдос» проф. Е.В.Луценко (отметим, что система «Эйдос» создана полностью с использованием только лицензионного инструментального программного обеспечения и на нее имеется 32 свидетельства РосПатента РФ);
- является одной из первых отечественных систем искусственного интеллекта персонального уровня, т.е. не требует от пользователя специальной подготовки в области технологий искусственного интеллекта: «имеет нулевой порог входа» (есть акт внедрения системы «Эйдос» 1987 года) (http://lc.kubagro.ru/aidos/aidos02/PR-4.htm);
- реально работает, обеспечивает устойчивое выявление в сопоставимой форме силы и направления причинно-следственных зависимостей в неполных зашумленных взаимозависимых (нелинейных) данных очень большой размерности числовой и не числовой природы, измеряемых в различных типах шкал (номинальных, порядковых и числовых) и в различных единицах измерения (т.е. не предъявляет жестких требований к данным, которые невозможно выполнить, а обрабатывает те данные, которые есть);
- имеет «нулевой порог входа», содержит большое количество локальных (поставляемых с инсталляцией) и облачных учебных и научных Эйдос-приложений (в настоящее время их 31 и более 362, соответственно:

http://lc.kubagro.ru/Source_data_applications/WebAppls.htm)
(http://lc.kubagro.ru/aidos/Presentation_Aidos-online.pdf);

- поддерживает on-line среду накопления знаний и обмена ими,

широко используется во всем мире (http://lc.kubagro.ru/map5.php);



- обеспечивает мультиязычную поддержку интерфейса на 51 языке. Языковые базы входят в инсталляцию и могут пополняться в автоматическом режиме;
- наиболее трудоемкие в вычислительном отношении операции синтеза моделей и распознавания реализует с помощью графического процессора (GPU), что на некоторых задачах обеспечивает ускорение решение этих задач в несколько тысяч раз, что реально обеспечивает интеллектуальную обработку больших данных, большой информации и больших знаний (графический процессор должен быть на чипсете NVIDIA);
- обеспечивает преобразование исходных эмпирических данных в информацию, а ее в знания и решение с использованием этих знаний задач классификации, поддержки принятия решений и исследования предметной области путем исследования ее системно-когнитивной модели, генерируя при этом очень большое количество табличных и графических выходных форм (развития когнитивная графика), у многих из которых нет никаких аналогов в других системах (примеры форм можно посмотреть в работе: http://lc.kubagro.ru/aidos/aidos18 LLS/aidos18 LLS.pdf);
- хорошо имитирует человеческий стиль мышления: дает результаты анализа, понятные экспертам на основе их опыта, интуиции и профессиональной компетенции;
- вместо того, чтобы предъявлять к исходным данным практически неосуществимые требования (вроде нормальности распределения,

абсолютной точности и полных повторностей всех сочетаний значений факторов полной независимости И аддитивности) их автоматизированный (АСК-анализ) системно-когнитивный анализ предлагает без какой-либо предварительной обработки осмыслить эти данные и тем самым преобразовать их в информацию, а затем преобразовать эту информацию в знания путем ее применения для достижения целей (т.е. для управления) и решения задач классификации, принятия решений поддержки и содержательного эмпирического исследования моделируемой предметной области.

В чем сила подхода, реализованного в системе Эйдос? В том, что она реализует подход, эффективность которого не зависит от того, что мы думаем о предметной области и думаем ли вообще. Она формирует модели непосредственно на основе эмпирических данных, а не на основе наших представлений о механизмах реализации закономерностей в этих данных. Именно поэтому Эйдос-модели эффективны даже если наши представления о предметной области ошибочны или вообще отсутствуют.

В этом и слабость этого подхода, реализованного в системе Эйдос. Модели системы Эйдос - это феноменологические модели, отражающие эмпирические закономерности в фактах обучающей выборки, т.е. они не отражают причинно-следственного механизма детерминации, а только сам факт и характер детерминации. Содержательное объяснение этих эмпирических закономерностей формулируется уже экспертами на теоретическом уровне познания в содержательных научных законах 14.

В разработке системы «Эйдос» были следующие этапы:

1-й этап, «подготовительный»: 1979-1992 годы. Математическая модель системы "Эйдос" разработана в 1979 и впервые прошла экспериментальную апробацию в 1981 году (первый расчет на компьютере на основе модели). С 1981 по 1992 система "Эйдос" неоднократно реализовалась на платформе Wang (на компьютерах Wang-2200C). В 1987 году впервые получен акт внедрения на одну из ранних версий системы «Эйдос», реализованную в среде персональной технологической системы «Вега-М» разработки автора (см. 2-й акт ПО ссылке: http://lc.kubagro.ru/aidos/aidos02/PR-4.htm).

2-й этап, «эра IBM PC и MS DOS»: 1992-2012 годы. Для IBМ-совместимых персональных компьютеров система "Эйдос" впервые реализована на языках CLIPPER-87 и CLIPPER-5.01 (5.02) в 1992 году, а в 1994 году уже были получены <u>свидетельства РосПатента</u>, первые в Краснодарском крае и, возможно, в России на системы искусственного интеллекта (слева приведена титульная видеограмма финальной DOS-версии системы «Эйдос-12.5», июнь 2012 года). С тех пор и до настоящего времени система непрерывно совершенствуется на IBM PC.

¹⁴ Ссылка на это краткое описание системы «Эйдос» на английском языке: http://lc.kubagro.ru/aidos/The_Eidos_en.htm

3-й этап, «эра MS Windows хр, 8, 7»: 2012-2020 годы. С июня 2012 по 14.12.2020 система «Эйдос» развивалась на языке Аляска-1.9 + Экспресс++ + библиотека для работы с Internet xb2net. Система «Эйдос-X1.9» хорошо работала на всех версиях MS Windows кроме Windows-10, которая требовала специальной настройки. Наиболее трудоемкие в вычислительном отношении операции синтеза моделей и распознавания реализует с помощью графического процессора (GPU), что на некоторых задачах обеспечивает ускорение решение этих задач в несколько тысяч раз, что реально обеспечивает интеллектуальную обработку больших данных, большой информации и больших знаний (графический процессор должен быть на чипсете NVIDIA).

4-й этап, «эра MS Windows-10»: 2020-2021 годы. С 13.12.2020 года по настоящее время система «Эйдос» развивается на языке <u>Аляска-2.0</u> + <u>Экспресс++</u>. Библиотека хb2net в ней больше не используется, т.к. все возможности работы с Internet входят в <u>базовые возможности языка программирования</u>.

5-й этап, «эра Больших данных, информации и знаний»: с 2022 года по настоящее время. С 2022 года автор и разработчик системы «Эйдос» проф.Е.В.Луценко вплотную занялся разработкой профессиональной версии системы «Эйдос» на языке Аляска+Экспресс, обеспечивающей обработку больших данных, информации и знаний (Big Data, Big Information, Big Knowledge) с использованием ADS (Advantage Database Server), а также на языке C# (Visual Studio | C#).

На рисунке 1 приведена титульная видеограмма DOS-версии системы «Эйдос» (годы эксплуатации: 1992-2012), а на рисунке 2 — текущей версии системы «Эйдос» (годы эксплантации: 2012- по настоящее время):



Рисунок 4. Титульная видеограмма DOS-версии системы «Эйдос» (до 2012 года)¹⁵

.

¹⁵ http://lc.kubagro.ru/pic/aidos titul.jpg

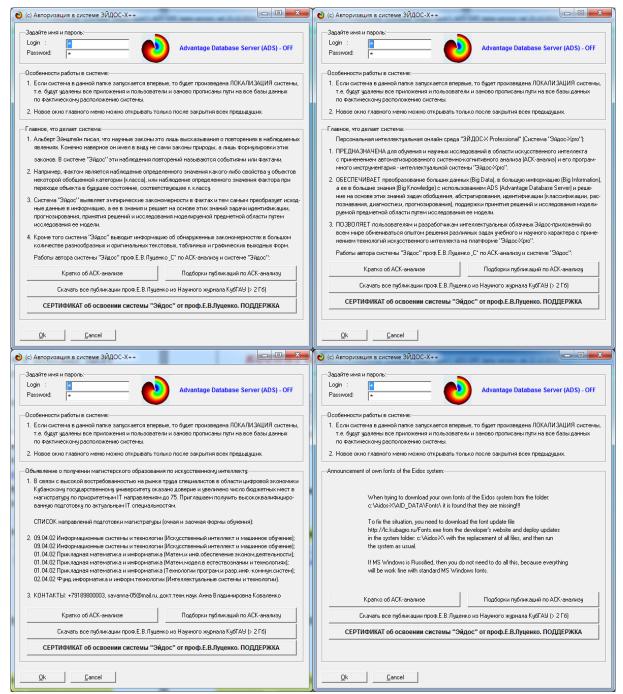


Рисунок 5. Титульные видеограммы текущей версии системы «Эйдос»

2.4. Задачи работы

Целью работы является решение поставленной проблемы.

Как уже показано выше, для работы с лингвистическими переменными целесообразно применить лингвистический АСК-анализ [22].

Достижение поставленной цели в АСК-анализе обеспечивается решением следующих *задач* и подзадач, которые являются *этапами* достижения цели:

<u>Задача-1.</u> Когнитивная структуризация предметной области.

<u>Задача-2.</u> Формализация предметной области.

<u>Задача-3.</u> Синтез статистических и системно-когнитивных моделей. Многопараметрическая типизация и частные критерии знаний.

<u>Задача-4.</u> Верификация моделей.

Задача-5. Выбор наиболее достоверной модели.

<u>Задача-6.</u> Системная идентификация и прогнозирование. Интегральные критерии знаний.

<u>Задача-7.</u> Поддержка принятия решений (Упрощенный вариант принятия решений как обратная задача прогнозирования, позитивный и негативный информационные портреты классов, SWOT-анализ; Развитый алгоритм принятия решений в АСК-анализе).

<u>Задача-8</u> исследование объекта моделирования путем исследования его модели, <u>включает ряд подзадач:</u>

- 1) инвертированные SWOT-диаграммы значений описательных шкал (семантические потенциалы);
 - 2) кластерно-конструктивный анализ классов;
 - 3) кластерно-конструктивный анализ значений описательных шкал;
 - 4) модель знаний системы «Эйдос» и нелокальные нейроны;
 - 5) нелокальная нейронная сеть;
 - 6) 3d-интегральные когнитивные карты;
- 7) 2d-интегральные когнитивные карты содержательного сравнения классов (опосредованные нечеткие правдоподобные рассуждения);
- 8) 2d-интегральные когнитивные карты содержательного сравнения значений факторов (опосредованные нечеткие правдоподобные рассуждения);
 - 9) когнитивные функции;
 - 10) значимость описательных шкал и их градаций;
- 11) степень детерминированности классов и классификационных шкал).

На рисунке 2 приведена последовательность преобразования исходных данных в информацию, а ее в знания и применения этих знаний для решения различных задач в системе «Эйдос»:

3. РЕЗУЛЬТАТЫ

3.1. Когнитивная ветеринария - ветеринария цифрового общества: дефиниция базовых понятий

Есть много мнений по вопросу о том, в каком обществе мы живем в начале XXI века. Ранее считали, что это постиндустриальное общество. Было даже такое мнение, что это будет общество развитого социализма или даже коммунистическое общество. Потом мнение по этому поводу изменилось. Сначала современное общество стали называть информационным обществом, а затем обществом, основанным на знаниях. Последняя новость в этой области состоит в том, что, повидимому, современное общество — это цифровое общество, т.е. общество,

основанное на цифровых технологиях, цифровых коммуникациях, цифровых технологиях обработки и передачи информации, и цифровых технологиях искусственного интеллекта. В цифровом обществе и науки должны перейти к цифровым интеллектуальным технологиям исследования. В частности возникает вопрос о том, не должна ли и ветеринария в цифровом обществе стать когнитивной ветеринарией. Развернутому и аргументированному (по мнению авторов) ответу на этот вопрос и посвящена данная работа. Методология и терминология в данной новой области еще не устоялась и не является общепринятой. Поэтому в данной работе много внимания уделено логике и методологии научного познания, терминологическим вопросам и дефинициям понятий.

Раздел 3.1 основан на работе [1]. Ссылки на литературные источники в данном разделе даны по списку литературы работы [1].

3.1.1. Введение

Есть много мнений по вопросу о том, в каком обществе мы живем в начале XXI века. Ранее считали, что это постиндустриальное общество. Было даже такое мнение, что это будет общество развитого социализма или даже коммунистическое общество. Потом мнение по этому поводу изменилось.

Сначала современное общество стали называть информационным обществом, а затем обществом, основанным на знаниях. Последняя новость в этой области состоит в том, что, по-видимому, современное общество — это цифровое общество, т.е. общество, основанное на цифровых технологиях: цифровых вычислительных ресурсах, цифровых коммуникациях, цифровых технологиях обработки и передачи информации и цифровых технологиях искусственного интеллекта.

В цифровом обществе и науки должны перейти к цифровым интеллектуальным технологиям познания.

В частности возникает вопрос о том, не должна ли и ветеринария в цифровом обществе стать когнитивной ветеринарией. Развернутому и аргументированному (по мнению авторов) ответу на этот вопрос и посвящена данная работа.

Научная терминология в данной новой области еще не устоялась и не является общепринятой. Поэтому в данной работе много внимания уделено терминологическим вопросам и дефинициям понятий.

Казалось бы тема данной работы хорошо разработана в научной и учебно-методической литературе. Чтобы убедиться в этом достаточно сделать следующие запросы в любой поисковой системе:

- «базы данных в ветеринарии»;
- «информационные системы в ветеринарии».

По этим запросам находятся без преувеличения тысячи сайтов и научно-методических работ. Среди них есть и очень хорошие, четкие и содержательные, хорошо на высоком научно-методическом уровне излагающие данную проблематику. Есть и много сайтов, содержание которых не обладает никакой новизной, выражаясь современным языком:

«с низкой оригинальностью», которые и не скрывают того, что их содержание просто скопировано с сайтов-первоисточников, на которые иногда дается ссылка, но довольно часто и не дается. Поэтому едва ли имеет смысл повторять содержание этих многочисленных сайтов, даже и «своими словами», в данной статье.

Однако при более внимательном ознакомлении с содержанием этих сайтов и научно-методических работ, на которых они основаны, возникает впечатление, что их авторы и составители часто не проводят особых различий между базами данных и информационными системами, по сути, допуская смешение содержания, размывание смысла этих понятий и на практике рассматривая их как синонимы. Традиционные определения этих понятий не выдерживают критики. Не дается четких определений этих понятий, не предлагаются критерии, позволяющие их разграничить. Не рассматривается сущность и содержание системного эффекта информационных систем, отличающего их от баз данных.

Авторы считают, что если в бытовой сфере это еще более-менее приемлемо, хотя и нежелательно, то в науке и образовании это уже недопустимо, особенно в наше время, когда уже созданы все предпосылки для упорядочения понятийного и терминологического аппарата и инструментария в этой предметной области. Тем более это совершенно недопустимо в такой науке как ветеринария, в которой традиционно совершено обоснованно придавалось очень большое значение точному определению содержания научных терминов и понятий, при котором не допускаются их разночтения и двусмысленные толкования (дефиниции понятий).

Таким образом, налицо *проблема*, которая традиционно рассматривается как несоответствие фактического и желаемого (целевого) положения вещей, противоречие между ними.

Проблема, решаемая в данной работе, стоит в том, что на практике допускается *смешение содержания*, *размывание смысла* понятий «данные», «информация», «знания», что в настоящее время является *недопустимым*, особенно в ветеринарии, и в этой связи необходима четкая дефиниция понятий.

Авторы имеют свое видение подходов к решению этой проблемы, которое и считают целесообразным рассмотреть подробнее в данной работе.

3.1.2. Данные - информация - знания

3.1.2.1. Традиционное определение понятия данных и его критика

Традиционно понятие данных определяется следующим образом: <u>данные</u> — это информация, записанная на каком-либо носителе (или находящаяся в каналах связи) и представленная на каком-то языке или в системе кодирования.

Это определение является общепринятым¹⁶, его легко найти во всех учебниках по информатике и информационным технологиям, однако оно не выдерживает никакой критики.

<u>Во-первых</u>, обычно определение понятия дается через более общее понятие и выделение специфического признака.

Например: млекопитающее – это животное (более общее понятие), выкармливающее своих детенышей молоком (специфический признак).

Если следовать этой логике, то понятие информации должно быть более общим, чем понятие данных, как и считается в традиционном определении, но в действительности все обстоит как раз наоборот. Кроме того, специфическим признаком информации, которая является данными, в традиционном определении оказывается то, что она записана на каком-то носителе на каком-то языке, в то время как и данные и информация, всегда записаны на каком-либо носителе в какой-то системе кодирования и невозможно даже представить себе информации, не записанной на носителе и не представленной на каком-либо языке.

<u>Во-вторых</u>, естественно, и более общее понятие, и специфический признак, должны быть *известны* и сами не требовать определения, иначе получится, что мы определяем одно неизвестное через другое неизвестное, иногда даже более неизвестное, чем первое. Но понятие информации не является более известным, чем определяемой через него понятие данных.

Например, представьте себе, что мы пытаемся дать инопланетянам, только-только исследовать которые начали нашу цивилизацию, представление о том, что такое бутерброд. Мы определяем для них что такое бутерброд и говорим: «бутерброд это хлеб, на который намазано масло». Но инопланетяне не знают не только того, что такое бутерброд, но и того, что такое хлеб и масло. И чтобы понять наше определение они уже немного волнуясь просят нас определить и эти понятия. Когда они спрашивают, что же такое хлеб, мы отвечаем: «Но так это же элементарно: это то, на что намазывают масло, когда делают бутерброд». И когда, наконец, они уже предчувствуя недоброе спрашивают, что же такое масло, мы отвечаем с ироничной улыбкой: «Но это Вы уже и сами должны были догадаться, – это то, что намазывают на хлеб, когда делают бутерброд». Мы уже не говорим о смысле слова: «намазывают». Все вместе взятые эти «определения» выглядят уже просто как издевательство над инопланетянами.

Наверное приведенный пример был бы даже смешен, если бы не был грустным, т. к. в науке подобный способ давать определения, как это ни странно, довольно распространен. Например нетрудно найти подобные «определения» материи и сознания друг через друга: материя — это то,

•

¹⁶ См., например: http://dic.academic.ru/dic.nsf/ruwiki/71919

что существует вне и независимо от сознания, а сознание это способность мозга отражать материю 17 .

Исходные данные об объекте управления обычно представлены в форме баз данных, чаще всего временных рядов, т. е. данных, привязанных времени. В соответствии методологией технологией c И автоматизированного системно-когнитивного анализа (АСК-анализ), развиваемой проф. Е. В. Луценко, для управления и принятия решений использовать непосредственно исходные данные не представляется возможным. Точнее сделать это можно, но результат управления при таком подходе оказывается мало чем отличающимся от случайного. Для реального же решения задачи управления необходимо предварительно преобразовать данные в информацию, о том, какие воздействия на объект моделирования к каким последствиям приводят (определить силу и влияния значений факторов на поведение управления), и в знания о том, какие значения факторов применить для воздействия на объект моделирования, чтобы он перешел в заранее заданные желательные целевые состояния.

3.1.2.2. Понятия данных, информации и знаний, сходство и различия между ними

3.1.2.2.1. Данные, подходы к определению

Продемонстрируем, что традиционное определение данных является ложным и абсурдным с логической точки зрения, применяя логический метод «Ложного основания». Это метод говорит о том, что если логическое следствие из некоторых исходных положений, является ложным, то и сами эти исходные положения также являются ложными.

Для этого попробуем дать определение понятия информации, основываясь на традиционном определении понятия данных и используя традиционный подход к научным определениям (дефинициям) через подведение определяемого понятия или термина под более общее понятие и выделение одного или нескольких специфических признаков.

При этом будем считать, как это, по сути, принято в традиционном определении (как мы это видели выше), что информация — это более общее понятие, чем данные, а специфическим признаком данных является то, что они записаны на носителе на определенном языке (это и есть ложное основание).

Эта попытка аналогична попытке дать определение более общего понятия «животное» на основе частного понятия «млекопитающее».

¹⁷ См., например: http://nounivers.narod.ru/bibl/diam9.htm: «Материя есть объективная реальность, существующая вне и независимо от сознания, тогда как сознание производно от материи и зависит от неё. Сознание есть отражение объективного мира в мозгу человека. Сознание-свойство высокоорганизованной материи, способность нашего мозга отражать вне нас существующий материальный мир.»

Исходное определение частного понятия: млекопитающее – это животное (более общее понятие), выкармливающее своих детенышей молоком (специфический признак).

Животное – это такое млекопитающее, которое:

- выкармливает своих детенышей молоком;
- или выкармливает своих детенышей не молоком.

По сути это означает, что животные это млекопитающие, но не только млекопитающие.

Информация – это такие данные, которые:

```
записаны на носителе на определенном языке, или записаны на носителе не на определенном языке; или не записаны на носителе на определенном языке, или не записаны на носителе не на определенном языке.
```

Таким образом на основе традиционного определения понятия «данные» *мы приходим к явно абсурдному результату*, т.к. информацию или данные, не записанные на каком-либо носителе на каком-либо языке или системе кодирования невозможно представить даже теоретически. Это и означает, что исходное традиционное определение данных, однозначным логическим следствием из которого является этот абсурдный результат, является таким же абсурдным, как и следствие из него, т.е. является ложным, неверным, что и т.д.

Следовательно, традиционное определение понятия данных является некорректным.

Но как же тогда корректно определить это понятие?

Однако, сделать это очень не просто по ряду причин:

Во-первых, по-видимому, понятие данных относится к числу наиболее общих понятий, выработанных человечеством. Это ясно из того, что о чем бы мы не рассуждали или не рассуждали, о каких бы объектах, процессах и явлениях внешнего и внутреннего мира, о природе обществе и человеке, мы все равно можем делать это только основываясь на каких-то данных об объекте рассуждения или познания. Поэтому возникают принципиально неразрешимая проблема поиска более общего понятия, чем понятие данных. Аналогичная неразрешимая проблема возникает при попытке определить традиционным путем (подведением под более общее понятие и выделение специфических признаков) другие предельно общие понятия, такие как бытие и небытие, материя и сознание, Бог, Вселенная и т.п. Впрочем подобные понятия можно пересчитать по пальцам одной руки.

<u>Во-вторых</u>, но даже если бы такое более общее понятие, чем понятие данных, удалось найти, все равно возникла бы проблема выделения таких специфических признаков, которые в этом более общем понятии позволяют выделить подмножество, соответствующее определяемому понятию: «Данные».

Поэтому в нашем распоряжении остается один вариант: описывать различные конкретные примеры данных всеми возможными признаками, а затем на основе этих описаний сформировать обобщенный образ данных, включив в него, признаки, вероятность наблюдения которых в данных намного выше, или на много ниже, чем в других понятиях, и исключив из него признаки, вероятность встречи которых в данных мало отличается от средней вероятности их встречи по всем формируемым обобщенным понятиям.

Такой ход рассуждений называется абдукция. Например: «Сократ смертен, Сократ человек, следовательно человек смертен». Таким образом мы что-то узнали об обобщенной категории «Человек». Понятно, что такой ход рассуждений является *правдоподобным*, но не гарантирует истинного результата, хотя в данном примере результат и получился истинный. В отличие от этого рассуждение от общего к частному всегда дает *истинный* результат, например: «Люди смертны, Сократ — человек, следовательно Сократ смертен». Отметим, что степень правдоподобности результатов абдукции возрастает, при увеличении числа примеров объектов, относящихся к различным обобщенным образам, и увеличении числа признаков, описывающих эти примеры.

Этот подход позволяет дать описательное или операциональное определение данных.

Однако этот подход предполагает рассмотрение не только понятия «Данные», но и других связанных с ним понятий, таких как «Информация» и «Знания», тем более что многие, вообще не видят между ними особых различий. Поэтому даже операциональное определение данных можно дать только в сопоставлении его с понятием информация, что мы и сделаем в следующем разделе.

3.1.2.2.2. Информация и данные

Информация определяется как осмысленные данные.

Данные — это более общее понятие, чем информация. Информация тоже является данными, но не каким угодно, а только осмысленными. Следовательно данные, — это такая информация, которая может быть как осмысленной, так и не осмысленной.

Это корректная попытка дать определение более общего понятия «Данные», через частный случай данных: понятие «Информация» и специфический признак информации: *осмысленность*.

Смысл данных, в соответствии с концепцией смысла Шенка-Абельсона [18], состоит в том, что известны причинно-следственные зависимости между событиями, которые описываются этими данными. Понятие причинно-следственных связей относится к реальной области. Данные же являются лишь моделью, с определенной степенью адекватности *отражающей* реальную предметной область. Поэтому в данных никаких причинно-следственных связей нет и выявить их в данных невозможно.

Но причинно следственные связи вполне возможно выявить между событиями, отражаемыми этими данными. Но для этого нужно предварительно преобразовать базу исходных данных в базу событий. Операция выявления причинно-следственных связей между событиями, отраженными в данных, называется «Анализ данных». По сути, анализ данных представляет собой их осмысление и преобразование в информацию.

Например, анализируя временные ряды, отражающие события на фондовом рынке, мы начинаем замечать, что если вырос спрос на какуюлибо валюту, то за этим обычно следует повышение ее курса.

Анализ данных включает следующие этапы:

- 1. Выявление событий в данных:
- разработка классификационных и описательных шкал и градаций;
- преобразование исходных в базу событий эвентологическую базу, путем кодирования исходных данных с применением классификационных и описательных шкал и градаций, т. е. по сути путем нормализации исходных данных.
- 2. Выявление причинно-следственных зависимостей между событиями в эвентологической базе данных.

В случае систем управления, событиями в данных являются совпадения определенных значений входных факторов и выходных параметров объекта управления, т. е. по сути, случаи перехода объекта управления в определенные будущие состояния, соответствующие классам, под действием определенных сочетаний значений управляющих факторов. Качественные значения входных факторов и выходных параметров естественно формализовать в форме лингвистических переменных. Если же входные факторы и выходные параметры являются числовыми, то их значения измеряются с некоторой погрешностью и фактически представляют собой интервальные числовые значения, которые также могут быть представлены или формализованы в форме порядковых лингвистических переменных (типа: «малые», «средние», «большие» значения показателей).

Какие же *математические меры* могут быть использованы для количественного измерения силы и направления причинно-следственных зависимостей?

Наиболее очевидным ответом на этот вопрос, который обычно первым всем приходит на ум, является: «Корреляция». Однако, в статистике хорошо известно, что это совершенно не так, т. к. для выявления причинно-следственных связей в соответствии с методом научной индукции (Ф. Бэкон, Дж. Милль) необходимо сравнивать

результаты по крайней мере в двух группах, в одной из которых фактор действовал, а в другой нет.

<u>Например</u>, на плакате, выпущенном полицией ¹⁸, написано: «По статистике, порядка 7,5-8 % аварий в России ежегодно совершается по вине водителей, находящихся в состоянии алкогольного опьянения» ¹⁹. Все. Точка. Больше ничего не написано. Однако, чтобы понять, является ли состояние алкогольного опьянения фактором, увеличивающим риск совершения ДТП или его тяжесть, этой информации недостаточно. Для этого обязательно необходима также информация о том, сколько процентов аварий в России ежегодно совершается по вине трезвых водителей. Но эта информация не приводится, поэтому формально здесь возможно три варианта:

- 1) по вине трезвых водителей аварий совершается меньше, чем по вине пьяных;
- 2) по вине трезвых водителей аварий совершается столько же, сколько по вине пьяных;
- 3) по вине трезвых водителей аварий совершается больше, чем по вине пьяных.

Первый вариант содержит информацию о том, что опьянение — это фактор риска совершения ДТП, второй — что это никак не влияет на риск совершения ДТП, а третий — что опьянение уменьшает его. Конечно, все понимают, что в жизни реализуется 1-ф вариант. Но об этом ведь в данном плакате нет прямых статистических данных. Таким образом, знак разности этих процентов определяет направление влияния этого фактора, а модуль этой разности силу его влияния, что и используется как один из частных критериев знаний в АСК-анализе и системе «Эйдос» [14, 20, 21].

Для преобразования исходных данных в информацию необходимо не только выявить события в этих данных, но и найти причинно-следственные связи между этими событиями. В АСК-анализе предлагается 7 количественных мер причинно-следственных связей, основной из которых является семантическая мера целесообразности информации по А.Харкевичу. Все эти меры причинно-следственных связей основаны на *сравнении* условных вероятностей встречи различных значений факторов при переходе объекта моделирования в различные состояния с безусловной вероятностью их встречи по всей выборке.

3.1.2.2.3. Знания и информация

<u>Знания</u> — это информация, полезная для достижения целей, т. е. для управления.

Значит для преобразования информации в знания необходимо:

¹⁸ Автор такой плакат видел, когда проходил медосмотр перед получением прав нового образца.

См., например: https://cnev.ru/polezno/stati/osnovnye-prichiny-dtp-pyanstvo-za-rulem

- 1. Поставить цель (классифицировать будущие состояния моделируемого объекта на целевые и нежелательные в какой-то шкале, лучше всего в порядковой или числовой).
- 2. Оценить *полезность* информации для достижения этой цели (знак и силу влияния).

Второй пункт, по сути, выполнен при преобразовании данных в информацию. Поэтому остается выполнить только первый пункт, т. к. классифицировать будущие состояния объекта управления как желательные (целевые) и нежелательные.

Знания могут быть представлены в различных формах, характеризующихся различной степенью формализации:

- вообще неформализованные знания, т. е. знания в своей собственной форме, ноу-хау (мышление без вербализации есть медитация);
 - знания, формализованные в естественном вербальном языке;
- знания, формализованные в виде различных методик, схем, алгоритмов, планов, таблиц и отношений между ними (базы данных);
- знания в форме технологий, организационных, производственных, социально-экономических и политических структур;
- знания, формализованные в виде математических моделей и методов представления знаний в автоматизированных интеллектуальных системах (логическая, фреймовая, сетевая, продукционная, нейросетевая, нечеткая и другие).

Таким образом, для решения сформулированной проблемы необходимо осознанно и целенаправленно последовательно повышать степень формализации исходных данных до уровня, который позволяет ввести исходные данные в интеллектуальную систему, а затем:

- преобразовать исходные данные в информацию;
- преобразовать информацию в знания;
- использовать знания для решения задач управления, принятия решений и исследования предметной области.

Процесс преобразования данных в информацию, а ее в знания называется анализ. Основные его этапы приведены на рисунке 2.

В системе «Эйдос» этот процесс осуществляется в следующей последовательности (рисунок 3).

Основные публикации автора по вопросам выявления, представления и использования знаний [13, 17].

Из вышеизложенного можно сделать обоснованный вывод о том, что *АСК-анализ и система «Эйдос» обеспечивают движение познания от* эмпирических данных к информации, а от нее к знаниям. По сути, это движение от феноменологических моделей, описывающих явления внешне, к содержательным теоретическим моделям [17].

Появляется все больше сайтов, посвященных искусственному интеллекту, в открытом доступе появляются базы данных для машинного

обучения (UCI^{20} , Kaggle²¹ и другие) и даже on-line интеллектуальные приложения, совершенствуется и интерфейсы, применяемые в Internet.

В этом смысле показательно приобретение разработчиком одной из первых и наиболее популярный по сегодняшний день глобальных социальных сетей Facebook Марком Цукербергом фирмы Oculus, являющейся ведущим в мире разработчиком и производителем амуниции виртуальной реальности.

Однако учащиеся и ученые до сих пор практически не замечают, что действует существует открытая масштабируемая давно И уже интерактивная интеллектуальная on-line среда для обучения и научных исследований, основанная на автоматизированном системно-когнитивном анализе (АСК-анализ) программном инструментарии его интеллектуальной системе «Эйдос», а также сайте автора.

Соотношение между содержанием понятий: «Данные», «Информация» и «Знания» наглядно показаны на рисунке 1. На рисунке 2 приведена схема преобразования данных в информацию, а ее в знания и решения на этой основе ряда задач в АСК-анализе и интеллектуальной системе «Эйдос».

Ниже рассмотрим основные компоненты этой среды подробнее.

3.1.2.2.4. От больших данных к большой информации, а от нее к большим знаниям

Internet постепенно интеллектуализируется и превращается из нелокального хранилища больших данных (big data) в информационное пространство, содержащее осмысленные большие данные, т. е. «большую информацию» (great info), а затем в пространство знаний или «когнитивное пространство», в котором большая информация активно используется для достижения целей (управления) и тем самым превращается в «большие знания» (great knowledge).

3.1.2.3. Основные термины баз данных, информационных и интеллектуальных систем

<u>Банк данных</u> — это базы данных плюс система *управления базами* ∂ *анных* (СУБД) (стандартные термины). СУБД — это, по сути, *система управления* ∂ *анными*.

<u>Информационный банк</u> — это информационные базы плюс информационные системы (предлагается стандартизировать эти термины). Информационная система — это, по сути, *система управления информацией*.

-

http://archive.ics.uci.edu/ml/datasets.html

https://www.kaggle.com/datasets

<u>Банк знаний</u> — это базы знаний плюс интеллектуальные системы (стандартные термины). Интеллектуальная система — это, по сути, *система управления знаниями*.

Существует очевидная параллель между терминами и понятиями, связанными с данными, информацией и знаниями, наглядно представленная в таблице 1.

Таблица 7 – ПАРАЛЛЕЛЬ МЕЖДУ ПОНЯТИЯМИ И ТЕРМИНАМИ, КАСАЮЩИМИСЯ ДАННЫХ, ИНФОРМАЦИИ И ЗНАНИЙ

Объект	Субъект	Система
База данных (БД)	Система управления базами данных	Банк данных =
	(СУБД)	БД+СУБД
Информационная	Информационная (аналитическая)	Информационный
база (ИБ)	система (система управления	банк = ИБ+СУИБ
	информационными базами – СУИБ)	
База знаний (БЗ)	Интеллектуальная система (система	Банк знаний =
	управления базами знаний – СУБЗ)	Б3+СУБ3

Поэтому неверно говорить о базах данных, понимая под этим по существу банки данных, т.к. база данных — это просто данные на носителях, а банк данных включает кроме самой базы данных еще и программную систему управления этими базами данных.

Автор предлагает «узаконить», т.е. стандартизировать термины, отмеченные в таблице 2 красным цветом. Это позволит упорядочить все эти термины в единой стройной системе, построенной на основе соотношения содержания понятий «данные», «информация» и «знания».

Это актуально, т.к. в настоящее время существуют явная путаница в понятий, встречающая использовании этих даже В названиях соответствующих дисциплин: «Управление знаниями», «Интеллектуальные информационные системы», «Представление знаний в информационных системах». Например, дисциплина «Управление знаниями» является гуманитарной ней изучаются слабо формализованные, не основанные на применении автоматизированных интеллектуальных систем, этапы, формы и методы управления знаниями²². Вместе с тем название этой дисциплины явно соотносится с названием дисциплины «Управление данными». Интеллектуальные системы часто информационными некорректно называются интеллектуальными было бы системами, тем же успехом ИХ ОНЖОМ «Интеллектуальные СУБД», но лучше и правильнее было бы называть их

зарубежных компаний; менеджмент знаний: по управлении знаниями (knowledge management)

²² Типичные вопросы, изучаемые в этой дисциплине: стратегия управления знаниями предприятия; организационная культура в контексте управления знаниями; измерение интеллектуального капитала; корпоративные знания: как ими управлять; интеграция знаний предприятия; бизнес держится на знаниях, сам того не зная; новые программы корпоративного обучения в среде управления знаниями: опыт зарубежных компаний; менеджмент знаний: подход к внедрению; общепринятых заблуждений об

как предложено: «Системы управления базами знаний». Дисциплина структуры данных» «Алгоритмы соотносится дисциплиной «Представление знаний в информационных системах», хотя ясно, что они представляются не в информационных, а в интеллектуальных системах, т.к. в информационных системах представляется информация, а не знания, а знания представляются в интеллектуальных системах. В настоящее время дисциплина «Интеллектуальные информационные системы» по своему содержанию включает «Представление знаний в информационных системах», тогда как из вышеизложенного ясно, что они должны соотносится по своему содержанию также, как СУБД и «Модели баз данных» (в которых обычно подробно преподается лишь одна реляционная модель).

Отметим также, что если применить данное выше определение знаний к моделям, описываемым в дисциплине «Представление знаний в информационных системах», то обнаруживается, что иногда в ней описываются не модели баз знаний, а модели баз данных или информационные модели. В частности это видно на примере семантических сетей, которые, по сути, представляют собой не более чем инфологическую модель реляционной базы данных.

Дисциплины «Управление знаниями» и «Представление знаний в интеллектуальных системах» по сути, представляют собой две части одной дисциплины и должны отражать не способы управления знаниями различной степени формализации (как в настоящее время), а описание автоматизированных интеллектуальных систем и баз знаний.

Существует дисциплина: «Алгоритмы и структуры данных». Предлагается ввести аналогичные дисциплины: «Алгоритмы и информационные структуры» (в АСК-анализе — это формализация предметной области и синтез модели) и «Алгоритмы структурирования знаний» (по содержанию близко к когнитологии, инженерии знаний, представлению знаний)».

Конечно, интеллектуальные системы являются и информационными (аналитическими) системами, и системами управления базами данных (рисунок 1). Информационные (аналитические) системы, являются системами управления базами данных (рисунок 1). Но не всякая система управления базами данных является информационной (аналитической) системой, а лишь такая, в которой в результате анализа данных выявляется преобразуются в информацию. они И информационная (аналитическая) система является интеллектуальной, а лишь такая, в которой в результате постановки цели и решения задачи управления (т.е. достижения цели) информация преобразуется в знание. Поэтому об авторах образовательных стандартов, в которых предлагается вести дисциплину: «Информационные интеллектуальные системы» можно сказать, что они не вполне понимают, о чем говорят.

Факт наличия причинно-следственных зависимостей может быть установлен методом хи-квадрат, а ее вид — многофакторным анализом. Однако факторный анализ позволяет обрабатывать данные лишь очень небольших размерностей (по числу факторов) и предъявляет чрезвычайно жесткие требования к наличию полных повторностей всех вариантов сочетаний факторов в исходных данных (т.е. данные не должны быть фрагментарными), что на практике выполнить удается крайне редко.

Поэтому большой интерес представляют другие подходы, обеспечивающие в различных предметных областях, в частности в ветеринарии, применение информационных и когнитивных технологий для выявление силы и направления причинно-следственных зависимостей в эмпирических данных.

При этом будут возникать новые направления науки. Широкое применение математических методов в экономике привело к возникновению такого направления науки и специальности ВАК РФ, как 08.00.13 - Математические и инструментальные методы экономики». Аналогично, ветеринария, широко применяющая информационные и когнитивные технологии для выявление силы и направления причинноследственных зависимостей в эмпирических данных, станет когнитивной ветеринарией.

3.1.2.4. Информационные системы как аналитические системы

Выше мы подробно рассматривали процесс преобразования данных в информацию путем автоматизированного осмысления данных. Процедура преобразования данных в информацию называется «Анализ».

Анализ данных включает следующие этапы:

- 1. Выявление событий в данных:
- разработка классификационных и описательных шкал и градаций;
- преобразование исходных в базу событий эвентологическую базу, путем кодирования исходных данных с применением классификационных и описательных шкал и градаций, т. е. по сути путем нормализации исходных данных.
- 2. Выявление причинно-следственных зависимостей между событиями в эвентологической базе данных.

Системы, обеспечивающие анализ данных, есть все основания называть информационными, т.к. они обеспечивают работу с информацией, и *аналитическими*, т.к. информация получена в результате *анализа* данных.

Таким образом термины «Информационная система» и «Аналитическая система» являются синонимами, а термин «Информационно-аналитическая система» является хотя и корректным по смыслу, но избыточным (переопределенным) и поэтому нежелательным.

Приведем другие примеры подобных переопределенных терминов: бутерброд с маслом, CD-диск, диссертационное исследование, комплексный системный анализ.

Применение подобных сочетаний слов может указывать на то, что те, кто его допускает не вполне понимают, что говорят. В частности применение термина: «Информационно-аналитическая система» может указывать на то, что применяющие его не понимают, что информационные системы являются аналитическими, а аналитические информационными по определению.

3.1.2.5. Интеллектуальные системы как системы управления. Развитый алгоритм принятия решений с применением SWOT- и кластерно-конструктивного анализа, ФСА и метода Директ-костинг

Выше мы уже видели, что знания — это информация, полезная для достижения цели. Знания содержат путь, способ или технологию (в т.ч. ноу-хау) достижения цели. А управление определяется как деятельность по достижению цели. Таким образом, все системы управления должны содержать модель, в которой отражены знания о том, как достичь цели в той или иной ситуации.

Методы Функционально-стоимостного анализа и «Директ-костинг» общеизвестны и популярны. По своим идеям и принципам Функционально-стоимостной анализ и метод «Директ-костинг» очень сходны, если не сказать тождественны. С одной стороны эти идеи весьма разумны, хорошо обоснованы теоретически и доказали свою эффективность на практике.

Однако широкому применению этих методов препятствует сложность получения больших объемов детализированной технологической и финансово-экономической информации, а также необходимость ее тщательного исследования компетентными специалистами, хорошо и содержательно разбирающимися в предметной области.

В этом и состоит противоречие между желанием применить методы ФСА и «Директ-костинг» сложностью это сделать на практике. Это противоречие представляет собой реальную проблему и часто обескураживает и вызывает разочарование этими методами.

В данной работе предлагается простое и эффективное решение данной проблемы, хорошо обоснованное теоретически, оснащенное всем необходимым методическим и программным инструментарием и широко и успешно апробированное на практике.

Предлагаемое решение основано на двух простых идеях:

- 1) вместо сбора и проведения содержательного исследования большого объема технологической и финансово-экономической информации применить подходы, приятные в теории управления;
- 2) для создания системы автоматизированного управления натуральной и финансово-экономической эффективностью затрат применить автоматизированный системно-когнитивный анализ и его программный инструментарий интеллектуальную систему «Эйдос».

В названии специальности 08.00.05 — Экономика и управление народным хозяйством, есть такие слова: «управление предприятиями, отраслями, комплексами, инновациями». Использование термина «Управление» предполагает, что есть модель, отражающая силу и направление влияния факторов на объект управления, и есть управляющая система, принимающая решения на основе этой модели. Однако, как правило, не только в научных статьях, но и в диссертациях по этой специальности мы ничего этого не видим, а видим лишь финансово-экономические расчеты. Если даже управляющие факторы и упоминаются, то обычно вопрос об их силе и направлении влияния на объект управления не только не решается, но даже и не ставится. Соответственно обычно вообще не упоминается какой-либо способ определения этой силы и направления влияния.

В данной работе предлагается подход, основанный на теории управления, снимающий этот недостаток [22].

3.1.2.5.1. Общая структура интеллектуальной автоматизированной системы управления

В теории управления известно, что в состав системы управления входят: объект управления, управляющая система, управляющие факторы, воздействующие на объект управления, информация обратной связи о состоянии объекта управления (рисунок 3).

Управляющая система принимает решения о значениях управляющих факторов на основе модели объекта управления.

Проблема состоит как в разработке этой модели на основе эмпирических данных (это скорее научная проблема), так и в ее применении в режиме реального времени в составе управляющей системы для поддержки принятия управляющих решений (это практическая проблема).

Обе эти проблемы решаются на основе автоматизированного системно-когнитивного анализа (АСК-анализ) и реализующей его интеллектуальной системы «Эйдос» [22].

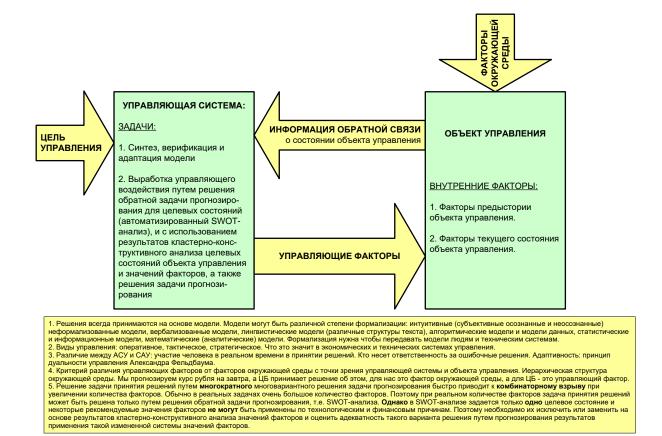


Рисунок 3. Принципиальная схема замкнутой адаптивной интеллектуальной автоматизированной системы управления

3.1.2.5.2. Алгоритм принятия управляющих решений в АСК-анализе и системе «Эйдос»

На рисунке 4 приведен алгоритм принятия управляющих решений в АСК-анализе и системе «Эйдос», использующий результаты SWOT-анализа классов и кластерного анализа значений факторов.

Алгоритм принятия решений

Традиционно, управляющие решения принимаются путем многократного решения задачи прогнозирования при различных значениях факторов и выбора такого их сочетания, управляющих обеспечивает перевод объекта управления в целевое состояние. Однако на реальные объекты управления действуют сотни и тысячи управляющих факторов, каждый из которых может иметь десятки значений. Полный перебор всех возможных сочетаний значений управляющих факторов приводит к необходимости решения задачи прогнозирования десятки и сотни тысяч и даже миллионы раз для принятия одного решения, и это является совершенно неприемлемым на практике. Поэтому необходим метод принятия решений не требующий значительных вычислительных ресурсов. Таким образом, налицо противоречие между фактическими и желаемым, в чем и состоит проблема, решаемая в работе. В данной работе предлагается развитый алгоритм принятия решений путем однократного решения обратной задачи прогнозирования (автоматизированный SWOTанализ), использующий результаты кластерно-конструктивного анализа целевых состояний объекта управления и значений факторов однократного решения задачи прогнозирования. Этим и обуславливается актуальность темы работы. Цель работы состоит в решении поставленной проблемы. Путем декомпозиции цели сформулированы следующие задачи, Когнитивно-целевая достижения являюшиеся этапами цели. структуризация предметной области; формализация предметной области (разработка классификационных и описательных шкал и градаций и формирование обучающей выборки); синтез, верификация и повышение достоверности модели объекта управления; прогнозирование, принятие решений и исследование объекта управления путем исследования его модели. В качестве метода решения поставленных задач применяется автоматизированный системно-когнитивный анализ и его программный инструментарий – интеллектуальная система «Эйдос». В результате работы предложен развитый алгоритм приятия решений, применимый в интеллектуальных системах управления. Основной вывод по результатам работы состоит в том, что предлагаемый подход позволил успешно решить поставленную проблему [18].

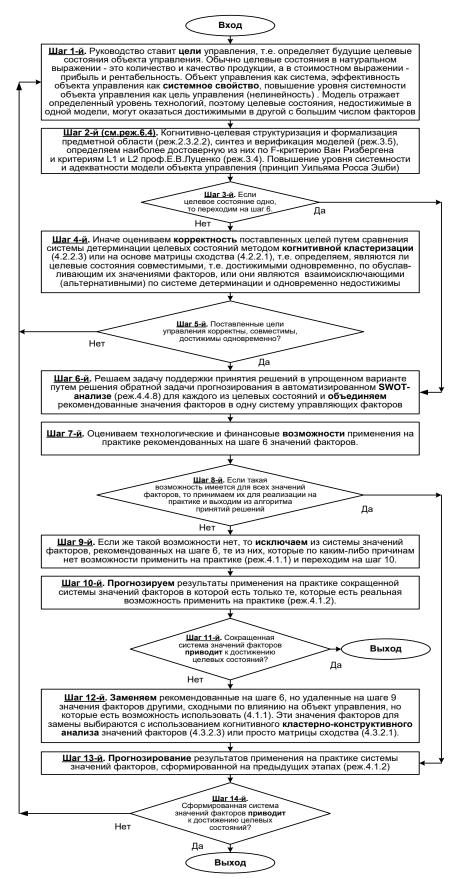
Предлагается следующий развитый алгоритм принятия решений в интеллектуальных системах управления на основе АСК-анализа и системы «Эйдос» (рисунок 2). Необходимо отметить, что система «Эйдос» обеспечивает решение всех задач, решение которых необходимо для реализации предлагаемого алгоритма: обратной задачи прогнозирования (автоматизированный SWOT-анализ) [19]; кластерно-конструктивный анализ целевых состояний объекта управления и значений факторов [20]; задачи прогнозирования [9-38].

Развитый алгоритм принятия решений АСК-анализа при его применении в интеллектуальных системах управления

<u>Шаг 1-й.</u> Ставим цели управления, т.е. определяем одно или несколько целевых состояний объекта управления. В натуральном выражении целевые состояния - это обычно количество и качество продукции, а в стоимостном выражении - прибыль и рентабельность ее производства и продажи.

<u>Шаг 2-й</u>. Проводим когнитивно-целевую структуризацию и формализацию предметной области, синтез и верификация статистических и системно-когнитивных моделей (СК-модели), определяем наиболее достоверную из них по F-критерию Ван Ризбергена и критериям L1 и L2 проф.Е.В.Луценко.

Шаг 3-й. Если целевое состояние одно, то переходим на шаг 6.



<u>Разработка автора</u>

Рисунок 4. Развитый алгоритм принятия управляющих решений в АСК-анализе и системе «Эйдос»

- <u>Шаг 4-й</u>. Иначе оцениваем корректность поставленных целей путем сравнения системы детерминации целевых состояний методом когнитивной кластеризации или просто на основе матрицы сходства, т.е. определяем, являются ли целевые состояния совместимыми, т.е. достижимыми одновременно, по обуславливающим их значениями факторов, или они являются взаимоисключающими (альтернатив-ными) по системе детерминации и одновременно достигнуты быть не могут.
- <u>Шаг 5-й</u>. Поставленные цели управления корректны, совместимы, достижимы одновременно?
- <u>Шаг 6-й</u>. Решаем задачу поддержки принятия решений в упрощенном варианте путем автоматизированного когнитивного SWOT-анализа целевых состояний.
- <u>Шаг 7-й</u>. Оцениваем технологические и финансовые возможности применения на практике рекомендованных на шаге 6 значений факторов.
- <u>Шаг 8-й</u>. Если такая возможность имеется для всех значений факторов, то принимаем их для реализации на практике и выходим из алгоритма принятий решений
- <u>Шаг 9-й.</u> Если же такой возможности нет, то исключаем из системы значений факторов, рекомендованных на шаге 6, те из них, которые по каким-либо причинам нет возможности применить на практике и переходим на следующий шаг.
- <u>Шаг 10-й</u>. Прогнозируем результаты применения на практике сокращенной системы значений факторов в которой есть только те, которые есть реальная возможность применить на практике.
- <u>Шаг 11-й</u>. Сокращенная система значений факторов приводит к достижению целевых состояний?
- <u>Шаг 12-й</u>. Заменяем рекомендованные на шаге 6, но удаленные на шаге 9 значения факторов другими, сходными по влиянию на объект управления, но такими, которые есть возможность использовать. Эти значения факторов для замены выбираются с использованием результатов когнитивного кластерно-конструктивного анализа значений факторов или просто матрицы сходства.
- <u>Шаг 13-й</u>. Прогнозирование результатов применения на практике системы значений факторов, сформированной на шаге 12.
- Шаг 14-й. Сформированная система значений факторов приводит к достижению целевых состояний? Если прогнозируемый результат применения на практике системы значений факторов, сформированной на шаге 12, по результатам прогнозирования приводит к переходу объекта управления в целевые состояния, то принимаем данную систему значений факторов для реализации на практике и выходим из алгоритма принятия решений. Если же прогноз показывает, что целевое состояние при использовании этой системы значений факторов не достигается, то задача управления не имеет решения в данной модели и осуществляется переход

на шаг 2 для качественного изменения модели с новыми исходными данными и расширенной системной значений факторов.

После выхода из алгоритма и реализации управляющих решений цикл управления, представленный на рисунке 2, повторяется. При этом результаты управления в любом случае, т.е. как при успешном достижении целевых состояний, так и в противном случае, учитываются в исходных данных для создания модели и осуществляется пересинтез модели. Поэтому непосредственно в процессе управления происходит постоянное улучшение качества интеллектуальной модели приятия решений путем ее самообучения с учетом фактических результатов управления. обеспечивается тем, что интеллектуальная система «Эйдос» является одновременно инструментом для синтеза и верификации моделей объекта управления, инструментом применения этих моделей для решения задач идентификации, прогнозирования, принятия решений и исследования моделируемой предметной области путём исследования ее модели. Достоверность созданных моделей оценивается с помощью F-меры Ван Ризбергена и ее мультиклассовых, нечетких обобщений, инвариантных относительно объема выборки (Луценко 2017). Система «Эйдос» не только обеспечивает решение этих задач, но и на данный момент, по-видимому, является единственной в мире системой, обеспечивающей решение всех этих задач на единой математической и технологической основе. При этом решение некоторых из этих задач по отдельности на данный момент автоматизировано «Эйдос», только системе например автоматизированный когнитивный SWOT-анализ, когнитивный кластернокогнитивных конструктивный анализ, построение когнитивных функций (Луценко 2017). Таким образом, развитый алгоритм принятия решений в интеллектуальных системах управления на основе АСК-анализа и реализуемый в системе «Эйдос», соответствует известному принципу дуального управления, предложенному в 50-х годах XX века в теории самонастраивающихся и самообучающихся систем замечательным советским ученым Александром Ароновичем Фельдбаумом.

3.1.2.5.3. Эксплуатация интеллектуальной АСУ в адаптивном режиме

Обратим внимание на то, что приведенный на рисунке 4 алгоритм принятия решений используется непосредственно в цикле управления (рисунок 3) и предусматривает постоянную адаптацию модели, а случае необходимости и ее пересинтез, что обеспечивает учет динамики моделируемой предметной области, т.е. как самого объекта управления, так и окружающей среды.

3.1.2.6. Понятие системы и системного эффекта (эмерджентности). Сущность и содержание системного эффекта

В данной работе мы постоянно говорим о различных видах систем:

- банках данных;
- информационных (аналитических) системах;
- интеллектуальных системах (системах искусственного интеллекта).

Поэтому целесообразно кратко теоретически рассмотреть, что такое вообще система, а затем что конкретно дает системный подход в понимании сущности и источников эффективности перечисленных видов систем.

Система есть совокупность взаимосвязанных элементов. Система имеет свойства, которых нет у ее элементов. Эти свойства называются системными или эмерджентными. Наличие у системы системных свойств обеспечивает ей преимущества в достижении цели или даже саму возможность достижения цели по сравнению с ее элементами. Системные свойства тем ярче выражены у системы, чем сильнее взаимосвязи между ее элементами. Элементы системы также являются системами, более низкого уровня иерархии, чем сама система. Кроме систем различных уровней иерархии во Вселенной ничего не существует. Все свойства всех объектов, процессов и явлений природы, общества и человека имеют системное происхождение и основаны на уровнях иерархии, этими свойствами не обладающих. Это касается даже таких свойств как существование и не существование, движение и пространство-время [14, 15, 16, 19].

Автор посвятил большое количество работ информационным мерам сложности, уровня системности и степени детерминированности систем, количественным подходам к их измерению с помощью предложенных коэффициентов эмерджентности [16]. Однако в данной работе подробно рассматривать эти работы нецелесообразно, т.к. они все находятся в полном открытом бесплатном доступе [16] https://www.researchgate.net/profile/Eugene_Lutsenko.

Таблица 2 – Состав и источник эффективности различных видов систем

Наименование вида	Элементы	Связи между	Системные (эмерджентные
системы	системы	элементами системы	свойства) системы
Банки данных	Базы данных (БД), системы управления базами данных (СУБД)	Даталогическая модель баз данных, инфологическая логическая модель баз данных (<i>отношения</i> между таблицами баз данных)	Возможность манипулирования данными, ведения баз данных, выполнения различных операций с данными в базах данных, поиск, физические и логические сортировка,
		————	фильтрация, группировки, выборки, формирование запросов и отчетов
Информационные	Информационные базы	Классификационные и	Выявление смысла данных и
(аналитические)	(ИБ), программное	описательные шкалы и	его использование для
системы	обеспечение	градации; база событий	решение различных задач,

	информационной	(эвентологическая	табличные и графические
	(аналитической)	база), базы причинно-	формы, отражающие смысл
	системы (системы	следственных	данных, т.е. причинно-
	управления	зависимостей между	следственные зависимости
	информационными	событиями в	между значениями факторов
	базами – СУИБ)	эвентологической базе	и результатами их влияния на
		данных	объект моделирования.
Интеллектуальные	Базы знаний (БЗ),	Сходство/различие	Решение задач
системы (системы	программное	между: 1) образами	идентификации,
искусственного	обеспечение	конкретных объектов и	прогнозирования,
интеллекта)	интеллектуальной	обобщенными	классификации, поддержки
	системы (система	образами классов; 2)	принятия решений по
	управления базами	образами классов, 3)	достижению поставленной
	знаний – СУБЗ)	значениями факторов;	цели, исследование
		4) количество	моделируемой предметной
		информации в	области путем исследования
		значениях факторов о	ее модели
		достижении цели	

Каждый последующий вид систем полностью включает элементы системы, связи между элементами и системные (эмерджентные свойства) всех предыдущих видов систем.

Целью создания банков данных; информационных (аналитических) систем и интеллектуальных систем (систем искусственного интеллекта) является не что иное, как получение системного (эмерджентного) эффекта.

3.1.2.7. Критерии идентификации банков данных, информационных и интеллектуальных систем

Эти критерии очевидны из предыдущего изложения.

Отметим, что довольно часто даже специалисты не проводят особых различий между базами данных и банками данных, называя банки данных базами данных. Обычно при этом подразумевается, что это не просто данные на носителях, но и программное обеспечение манипулирования ими. Таким образом, научная терминология используется неправильно, некорректно.

Аналогично часто банки данных, т.е. СУБД с базами данных, довольно часто называют информационными системами, подразумевая при этом, что в базах данных содержится информация, а не данные, т.е. не понимая, чем отличается информация от данных.

Более того, даже в учебниках по моделям представления знаний в базах знаний интеллектуальных систем часто фактически описываются не базы знаний, а информационные базы, или даже просто базы данных.

Поэтому вопрос о критериях идентификации банков данных, информационных и интеллектуальных систем является весьма актуальным. Если его решить то терминологическая путаница в головах учащихся и специалистов, а также в соответствующей литературе, может несколько уменьшится. Хотя если оставаться на реальной почве, то можно признать, что надежд на это немного.

Опираясь на таблицу 2 мы можем сказать, что если в программной системе, работающей с базами данных, есть классификационные и описательные шкалы и градации; база событий (эвентологическая база), причинно-следственных зависимостей событиями между эвентологической базе данных, т.е. система обеспечивает выявление смысла данных и его использование для решение различных задач, то есть все основания называть эту систему информационной или аналитической системой.

Опираясь на таблицу 2 мы можем сказать, что если в программной системе, работающей с базами данных, количественно выявляется сходство/различие между: 1) образами конкретных объектов обобщенными образами классов; 2) образами классов, 3) значениями факторов; 4) количество информации в значениях факторов о достижении используется для решения задач идентификации, прогнозирования, классификации, поддержки принятия решений по достижению поставленной цели, исследование моделируемой предметной области путем исследования ее модели, то есть все основания называть эту систему интеллектуальной системой или системой искусственного интеллекта (СИИ).

Надо признать, что эти требования довольно жесткие и многие из систем фактически не оправдывают своих названий, т.е. в названии системы заявляются более развитие функция, чем система поддерживает фактически.

С другой стороны вполне понятно желание разработчиков назвать свою систему красивым и модным сочетанием слов: «Информационная система», «Аналитическая система» или даже «система искусственного интеллекта», «Интеллектуальная система».

Но надо понимать, что когда разработчик делает это по сути не понимая смысла используемых им терминов, то он часто поддается соблазну выдать желаемое за действительное и вводит потенциальных пользователей в заблуждение, т.е. попросту обманывает их.

В этом нет ничего удивительного, специалисты по рекламе практически всегда так и делают²³, но это не может оправдать ни разработчиков программных систем, ни специалистов по рекламе.

3.1.2.8. Некоторые выводы и перспективы

Авторы считают, что по результатам данной работы можно обоснованно сделать следующие выводы:

²³ Например, показывают рекламный ролик с молодой женщиной, которой не более 25 лет, и вдруг она заявляет: «мне 40», и улыбается, а мы уже сами догадываемся, что а выглядит она так молодо потому, что использует рекламируемый ею чудодейственный крем, омолаживающий кожу лица.

- 1. В настоящее время ситуация с использованием научных терминов в области применения банков данных, информационных (аналитических) и интеллектуальных систем, баз данных, информационных баз и баз знаний нельзя признать удовлетворительной.
- 2. Предлагается система научных терминов, позволяющая упорядочить терминологию в этой предметной области.
- 3. Применение этой терминологии и критериев идентификации банков данных, информационных (аналитических) и интеллектуальных систем при преподавании информационных технологий и проведении научных исследований и разработок с применением этих технологий в различных предметных областях позволит повысить качество и адекватность принимаемых решения в этих областях.

В качестве перспективы предлагается ввести в номенклатуру научных специальностей ВАК РФ специальность: «Когнитивная ветеринария».

3.1.3. Этапы познания факт (конкретная онтология, фреймэкземпляр), эмпирическая закономерность, эмпирический закон, научный закон, философский закон

Процесс познания движется от эмпирического к теоретическому, от феноменологических моделей к содержательным, от формы к содержанию, от явления к сущности. От частного к общему и всеобщему, от познания локальных в пространстве-времени закономерностей, к познанию глобальных закономерностей. Принцип относительности и принцип аналогии.

В процессе познания современная наука использует различные формы познания: эмпирическое познание (чувственное познание, наблюдение, эксперимент), и интеллектуальную форму познания (логику и абстрактное мышление). Научные гении явно используют еще также и интуицию.

3.1.3.1. Факт (конкретная онтология, фрейм-экземпляр)

Понятие факта многими воспринимается как нечто элементарное, с чего начинается любое исследование. Однако это понятие далеко не так элементарно, как иногда кажется. Понятие факта связано с понятиями измерения и классификации, определения степени выраженности свойства объекта и определения принадлежности объекта к определенной группе.

<u>Факт далеко не элементарен и имеет довольно сложную структуру:</u>

1. Получение эмпирических данных об объекте исследования — это результаты наблюдения или эксперимента, полученные с помощью сенсоров или датчиков.

- 2. Формирование конкретного образа наблюдаемого объекта или его состояния на основе эмпирических данных.
- 3. Идентификация (классификация) конкретного образа наблюдаемого объекта или его состояния, т.е. определение его принадлежности к обобщающим категориям (классам).

Такую же структуру, как понятие факта, имеют также понятия конкретной онтологии (понятие из теория искусственного интеллекта) и фрейма-экземпляра (понятие из модели представления знаний Марвина Мински, 1975). По сути, есть все основания понятия факта, конкретной онтологии и фрейма-экземпляра рассматривать как синонимы

<u>На любом из этапов формирования факта могут быть различные</u> ошибки [3-5].

На этапе получения эмпирических данных об объекте исследования: результаты наблюдения или эксперимента наибольшую опасность представляет ошибка гипостазирования, суть которой состоит в том, что свойства способа наблюдения неоправданно переносятся на объект наблюдения. Например, в Изумрудном городе²⁴, при входе в который, как известно всем в обязательном порядке надевали зеленые очки, зеленый цвет очков ошибочно приписывался стекляшкам в дороге и в стенах домов, и даже облакам. Таким образом, объекту наблюдения приписывались свойства способа наблюдения. Нам это кажется чрезвычайно наивным. Однако это не так, т.к. мы ведем себя не намного лучше, всерьез считая, что светофор зеленый, тогда как немного поразмыслив мы понимаем, что зеленый цвет — это всего лишь субъективная форма в которой мы осознаем тот факт, что светофор излучает электромагнитные колебания такой длины волны (и частоты), которые мы осознаем как зеленый цвет.

На этапе формирования конкретного образа наблюдаемого объекта или его состояния на основе проблема может состоять в том, что довольно часто адекватный конкретный образ наблюдаемого объекта или явления не может быть создан из-за естественных ограничений способа наблюдения или способа обработки эмпирических данных. Например, мы привыкли к тому, что не видим крыльев пчелы, когда она зависает над цветком, не видим вращающихся лопастей винта самолета и даже не видим спиц в колесе едущего велосипеда. Вернее мы их видим, но неадекватно, т.е. не как крылья, не как лопасти и не как спицы, а как «нечто расплывчатое и полупрозрачное».

Ho, по-видимому, наиболее сложным интеллектуальным И котором осуществляется идентификация является 3-й этап на (классификация) конкретного образа наблюдаемого объекта или его определение его принадлежности к обобщающим т.е. категориям (классам). На этом этапе может возникнуть две проблемы.

 $^{^{24}}$ Волшебник изумрудного города — Волков А.М.

Первая проблема: по сути невозможно верно идентифицировать наблюдаемый явление, его конкретный объект или если некачественно сформирован или вообще не сформирован на предыдущем этапе, если это образ «чего-то расплывчатого и полупрозрачного». Вторая проблема: даже если конкретный образ является адекватным, то он все равно не может быть идентифицирован, если его по сути не с чем идентифицировать, т.е. в модели предметной области познающего субъекта отсутствуют обобщенные категории (классы), к которым действительно относится наблюдаемый объект или явление.

Интересно, что еще Ориген говорил, что *чудеса противоречат не* законам природы, а лишь нашим представлениям о законах природы. Поэтому если наблюдаемые факты противоречат законам природы, то значит у нас неверное представление об этих законах, неверная теория, которую надо развивать. Но так происходит далеко не всегда. Не зря ведь как бы в шутку говорят: «Если факты противоречат теории, то тем хуже для фактов». Наверное это было бы смешно, если бы не было грустно. Таким образом, описание факта всегда отражает не только саму наблюдаемую реальность, но и самого исследователя, прежде всего его уровень компетентности.

На эмпирическом этапе познания описания новых фактов всегда первоначально делается в старых понятиях, которыми владеют исследователи, и, поэтому, является неадекватным. И лишь на более позднем теоретическом этапе научного познания формируются новые адекватные понятия для описания новых фактов (К.Маркс). Таким образом описание факта содержит не только описание реальности, но всегда несет отпечаток времени и места, где жил его наблюдатель и присущих ему ограничений.

Сохранилось легенды папуасов Океании, описывающие прибытие каравелл европейцев. Они описывают, что в бухте неожиданно появилось несколько маленьких островов с крутыми берегами. На островах было по 2-3 высоких прямых дерева без листьев, опутанных паутиной. На этих островах также были существа, напоминавшие людей, но другой формы тела, достающие из тела и засовывающие обратно в него какие-то предметы. Эти существа ели сырое мясо крокодилов и выбрасывали их шкуру в море (арбузы), кроме того они дышали дымом. На кораблях были мачты и снасти, испанцы были в одежде с карманами, они ели арбузы и курили. Надо сказать, что описание, данное аборигенами, является довольно точным, но в тоже время очевидно, что они совершенно не поняли, что же они видели. Примерно так же и по аналогичным причинам люди описывали плоскую Землю, встающее утром из-за горизонта и садящееся вечером за горизонтом Солнце, метеориты, отколовшиеся от небесной сферы, примерно также и мы описываем НЛО.

Таким образом старые понятия могут и не быть адекватным инструментом для описания фактов. В этом случае описание фактов с их применением получится неадекватным и для их адекватного описания необходимо выработать адекватные для этой задачи «новые термины и понятия». Существует зависимость терминологической формулировки фактов от господствующей теории у тех исследователей, кто их выявляет. Это значит, что когда информация о фактах вербализуется, то по сути факты при этом интерпретируются, т.е. невольно теоретически позиций господствующей осмысливаются теории. Поэтому c формулировках фактов есть информация не только о том, что наблюдалось эмпирически, но и том, как эти наблюдения поняли наблюдатели.

Но от наблюдателя зависит не только описание фактов, но в ряде случаев и сами факты. Дело в том, что наблюдатель (непосредственно или с помощью измерительных приборов) получает информацию о факте в процессе взаимодействия с объектом познания, при этом не только объект познания влияет на него, но и он влияет на объект познания. В качестве примера можно привести эксперименты о зависимости физических свойств электронов, проявляющихся при их интерференции на щелях, от их наблюдения с помощью света (Комптон-эффекта)²⁵.

Факт – это не описание того, что было в действительности (как чаще всего ошибочно думают), а описание того, как наблюдатель или исследователь понял или не понял то, что было в действительности.

Факт (конкретная онтология, фрейм-экземпляр) – это соответствие общего и особенного, дискретного и континуального, экстенсионального и интенсионального [6].

Область действия факта локальна, т.е. факт описывает нечто, произошедшее в определенном месте и в определенное время, «здесь и сейчас».

3.1.3.2. Эмпирические закономерности

Это закономерности, обнаруженные одним исследователем в собранных им фактах. Область действия эмпирических закономерностей это сами факты, на которых они обнаружены (действие в узкой эмпирической области).

3.1.3.3. Эмпирические законы (феноменологические модели и модель «Черного ящика»)

Это эмпирическое закономерности, обнаруженные И подтвержденные многими исследователями на собранных ими фактах, работающими в разных местах и в разное время. На основе эмпирического исследования строится феноменологическая модель объекта исследования

См., например: http://vivovoco.astronet.ru/VV/Q_PROJECT/FEYNMAN/LECTURE6.HTM

(модель «черного ящика») и формулируется эмпирический закон. Область действия эмпирического закона шире, чем эмпирической закономерности и расширяется с фактически исследованной предметной области на всю область, которая в принципе может быть эмпирически исследована.

Иногда факты определяют как явления или их свойства, установленные эмпирически и допускающие проверку другими исследователями в другом месте и в другое время (верификацию) и успешно прошедшие такую проверку. Только тогда факты считаются существующими и достоверными.

По мнению автора, в это определение факта является некорректным по двум причинам:

- 1. Так как в начале этого определения говорится о том, что это определение фактов, а в конце о том, что оно касается фактов, существование которых установлено достоверно. В результате вообще непонятно что определяется в данном определении: просто факты или только достоверно установленные объективно существующие факты. Очевидно, что из этого следует необходимость доработки или уточнения данного определения.
- 2. В данном определении фактов активно используются развитые теоретические представления о принципе наблюдаемости и принципе относительности, которые сами описывают свойства фактов в зависимости от условий их наблюдения. Таким образом, в данном определении допущена логическая ошибка, которая называется «логический круг» или определение неизвестных понятий друг через друга (т.е. определение одного неизвестного через другое неизвестное), а не через другие известные более простые и более фундаментальные понятия. Пример такой логической ошибки мы видим при определении понятий «бутерброд», «хлеб», «масло». 1-е определение: «Бутерброд – это хлеб с маслом». 2-е определение: «Масло это то, что намазывают на хлеб, когда делают бутерброд». 3-е определение: «Хлеб – это то, на что намазывают масло, когда делают бутерброд». Нельзя не отметить, что примерно так же в философии определяются категории «Материя» и «Сознание»: «Материя - это то, что существует независимо от сознания», а «Сознание это то, что отражает материю».

Принцип наблюдаемости утверждает, что объективное существование может считаться установленным только для тех процессов и явлений, которые наблюдались несколькими, по крайней мере двумя, независимыми способами.

По мнению автора, *измерение* — это построение модели объекта на основе наблюдений и исследование этой модели для установления ненаблюдаемых характеристик объекта. Чем больше независимых способом (или датчиков) используется при измерении величины, чем тем

меньше погрешность измерения и выше достоверность с которой установлено ее существование и значение.

Например, заряд электрона получается одинаковым и при его определении путем исследования прохождения токов в электролитах, и путем наблюдения отклонений траекторий электронов в магнитном поле. Это является основание для утверждения об объективном существовании такого заряда.

Если же какое-то свойство наблюдается только одним способом, то невозможно понять, является ли это свойство свойством наблюдаемого объекта или оно наблюдается вследствие постоянной погрешности самого способа его наблюдения. Например, непонятно (пока не снимешь очки), действительно ли в Изумрудном городе полно изумрудов или просто стекляшки выглядят как изумруды из-за зеленых очков.

Принцип наблюдаемости играет важную методологическую роль не только в физике. В работе [16] тотальная ложь рассматривается как стратегическое информационное оружие общества периода глобализации и дополненной реальности. Рассматривается возможность применения в современном обществе принципа наблюдаемости, как общепринятого в физике критерия реальности. Показано, в каких случаях применение данного принципа в исследованиях общества приводит к общественным иллюзиям, а когда дает адекватные результаты. Предлагаются понятие: «Степень виртуализации общества» и количественная шкала для ее измерения, а также вводится понятие «Общественный умвельт» под которым понимается область общества, существенно отличающаяся от остальных своими фундаментальными закономерностями.

Таким образом по сути дела *принцип наблюдаемости дает* критерий достоверности факта, присвоить фактам объективный или субъективный статус [3, 16].

Принципы относительности утверждает, что *природа устроена* таким образом, что все физические явления и процессы протекают совершенно одинаково во всех системах отсчета, движущихся по инерции.

Из принцип относительности вытекает много следствий, на нем основаны фундаментальные физические теории (СТО и ОТО). Однако сам он не имеет какого-либо обоснования, кроме философско-эстетического предположения, что природа должна быть устроена просто и красиво (которое, по-видимому, правильно).

Мы отметим два следствия из принципа относительности.

<u>Во-первых</u>, если проводить эмпирические исследования в одном месте и в одно время, то результатами этих исследований можно воспользоваться в другом месте и в другое время. По сути, этот принцип позволяет расшить область действия фактов с локальной на глобальную, делокализовать факты. Во многом именно поэтому и имеет смысл проводить исследования.

<u>Во-вторых</u>, появляется невозможность определить, движется ли лаборатория или покоится, проводя любые физические эксперименты внутри нее. Если имеются в виду только механические эксперименты, то речь идет о принципе относительности Галилея, если кроме того и оптические (электромагнитные) – то о принципе относительности Галилея-Эйнштейна, если об информационных, психологических и социально-экономических процессах и явлениях – то о принципе относительности Луценко [17, 18].

Применимость принципов наблюдаемости и принципов относительности в социально-экономических явлениях.

работе [17] автором сделана попытка расширить область применения принципов, аналогичных принципам относительности Галилея Эйнштейна И рассматривать как важнейшие ИХ методологические принципы не только для физики, но и для других наук, т.е. как фундаментальные общенаучные методологические принципы. В этой связи необходимо отметить, что в некоторых других науках также были выработаны аналогичные по смыслу методологические принципы. Например в статистике и социально-экономических науках применяются понятия исследуемой и тестовой выборки и генеральной совокупности, рассматриваются вопросы репрезентативности исследуемой выборки по отношению к генеральной совокупности, а также вопросы такого исследуемой выборки, которое позволяет отражаемую ей генеральную совокупность. В теории временных рядов развито также расширение понятий «репрезентативная выборка» «генеральная совокупность» (которые ассоциируются c пространственными множество» **ПИКИТКНОП** «локальное И «делокализация») временную область. Если на закономерности, выявленные на каком-то отрезке временного ряда (т.е. исследуемая выборка) действуют для него и далее определенный период времени, то период называется «период эргодичности» (т.е. генеральная совокупность). Эргодичность процесса нарушается в точке бифуркации, в качественно изменяются и начинают действовать закономерности, ранее исследуемая выборка теряет репрезентативность и модели, созданные на ее основе, теряют адекватность.

Принцип аналогии ПО смыслу сходен принципами относительности и известен с глубокой древности. Достаточно сказать, что величайшему мудрецу древнего Египта Гермесу Трисмегисту («Трижды Величайший»), обожествленному впоследствии как бог приписываются такие слова, записанные на изумрудной скрижали²⁶: «Что вверху, то и внизу, каждая нить в великой ткани мироздания проходит от самого верха до самого низа».

существует легенда, согласно которой изумрудной скрижалью Гермеса Трисмегиста владел Александр Македонский, но знал ли он ее сокровенный смысл остается неизвестным.

В работе [17] предложен обобщенный принцип относительности, аналогичный принципу относительности Галилея-Эйнштейна, но для всех видов реальных и виртуальных систем, а не только физических, высказана гипотеза о его взаимосвязи с теоремой Эмми Нётер и законами сохранения энергии, импульса и момента импульса в социально-экономических и психологических системах. На основе информационной теории времени и информационной теории стоимости (Е.В. Луценко, 1979) сделаны выводы о неравномерности хода времени в социальных системах, неоднородности и анизотропности экономического пространства и нарушении законов сохранения энергии, импульса и момента импульса в социальных системах, и соответственно, о невыполнении или лишь частичном выполнении них обобщенного принципа ДЛЯ относительности. Предложены новые понятия физического и социально-экономического умвельта и с их использованием сформулированы физический и социально-экономический антропные принципы [3].

Итак, принципы наблюдаемости и относительности позволяют установить степень достоверности фактов и расшить область их применения, т.е. по сути, перейти от исследования эмпирических закономерностей к формулировке эмпирических законов, выраженных форме феноменологических моделей.

Феноменологические модели могут вполне адекватно отражать результат действия внешних факторов на процессы и явления, но при этом описывают процессы и явления внешне, не рассматривая их внутреннюю структуру, т.е. не рассматривая, каким образом внешние факторы влияют на внутреннюю структуру и каким образом изменения в этой внутренней структуре обусловливают изменение внешне наблюдаемых свойств этих процессов и явлений. В современной науке такие модели называются феноменологическими моделями, наиболее распространенное название которых «Модели черного ящика». Такие модели широко применяются в автоматической теории управления, в которой влияние внешних факторов на систему описывается передаточной функцией. Связанные с этой проблематикой вопросы рассматриваются в работах [19, 20].

Например, (несколько упрощая) можно сказать, что в сельском хозяйстве изучается влияние природных и технологических факторов, а также свойств сортов и пород, на количественные и качественные результаты производства сельхозпродукции, а также методы прогнозирования и поддержки принятия решений, направленные на достижение заданных результатов. А в биологических науках, а также биохимии, биофизике и т.п., изучаются механизмы влияния тех же самых факторов на эти результаты.

3.1.3.4. Научные законы (движение от феноменологических моделей к содержательным, от эмпирического к теоретическому познанию)

Дальнейшее движение процесса познания — это движение от феноменологических моделей к содержательным. Суть этого процесса состоит в том, что процесс познание переходит от познания чувственновоспринимаемой или познаваемой другими эмпирическими методами внешней стороны явлений и процессов к познанию их сущности. В отличие от внешней стороны сущность явлений и процессов при обычной наиболее массовой в настоящее время форме сознания не является непосредственно воспринимаемой и осознаваемой и для ее познания в настоящее время используется интеллектуальная форма познания: мышление и логика.

Задачей мышления является разработка такой теории изучаемых процессов и явлений, которая бы правильно объясняла эмпирически наблюдаемые их свойства. Таким образом, теория описывает некий «внутренний механизм» изучаемых процессов и явлений, объясняющий их внешне наблюдаемые свойства.

Разработка новой теории — это процесс многоэтапный. На первом этапе выдвигается научная гипотеза о причинах действия эмпирического закона. Если оказывается, что научная гипотеза имеет прогностическую силу, т.е. предсказывает новые наблюдаемые на опыте ранее неизвестные процессы и явления, то она приобретает статус научного закона.

Например, 100 лет назад Альберт Эйнштейн в рамках созданной им теории гравитации (ОТО) предсказал существование гравитационных волн, которые недавно были экспериментально обнаружены.

<u>Определение научного закона:</u> научный закон — это такой эмпирический закон, который действуют везде, где сохраняют силу и действуют причины его действия, описанные в теории, объясняющей причины и механизм действия данного эмпирического закона.

<u>Область действия</u> научного закона расширяется на неограниченную область не всегда доступную эмпирически, даже в принципе и в перспективе.

Научные законы получаются из эмпирических методом научной индукции (Ф. Бэкон, Дж. Милль) 27 :

- строится *содержательная* модель, «объясняющая», почему действует эмпирический закон;
- делается научное обобщение: эмпирические законы выполняются не только во всех исследованных случаях, но и во всех остальных, где сохраняется действие причин их выполнения.

3.1.3.5. Философское обобщение

Философский закон — это придание научному (или даже эмпирическому) закону статуса *всеобщности*, т.е. ничем неограниченное *расширение области действия* научных законов²⁸.

По мнению автора философское обобщение — это неоправданное и очень рискованное предельное обобщение, которое никогда не может быть в достаточной мере (для столь ответственного решения) аргументировано и обосновано, и, по глубокому убеждению автора, в конечном счете оно всегда ошибочно, что рано или поздно и выясняется (правда обычно поздно).

3.1.3.6. Перспективы применения научного метода к постановке и решению философских проблем и конец философии

История науки свидетельствует, что науки одновременно, определенной В последовательности a усложнения предмета их исследования: физика, химия, биология, науки о человеке и обществе. Понятно, что предметы исследования этих наук не только существовали и до их возникновения, но и исследовались до возникновения этих наук, но исследовались они не научными методами, а в рамках философии, которая выступала в роли своеобразной «преднауки», «прародительницы» и одновременно «матери всех наук»²⁹.

<u>Принцип</u> возникновения новых наук состоит в том, что они «отмежевываются» от преднауки философии когда начинают применять научный метод, так как именно этот метод позволял перевести знания из умозрительной формы в точно установленную доказательную форму.

При возникновении новых наук предмет философии сужается, так как часть предмета философии становится предметом изучения этих новых конкретных наук. Конкретные науки обеспечивают более глубокое, детальное, достоверное и доказательное изучение предмета познания, чем это было возможно в философии. В результате этого процесса в настоящее время предметом философии является лишь основной вопрос философии, а также диалектика, логика и теория познания.

По мнению автора «конец философии» наступит, когда научный метод будет применен к познанию не только материи (что уже сделано), но и сознания, а также к исследованию их *отношения*, т.е. к постановке и решению основного вопроса философии [3], и это станет реально возможным только при высших формах сознания, при которых людям станут доступными другие формы познания, в частности,

См., например: http://bookz.ru/authors/nina-bu4ilo/filosofi_921/page-4-filosofi_921.html

70

ни в пространстве, ни во времени, ни по предметной области, ни по уровню иерархического строения Вселенной как системы, ни каким-либо другим образом

интеллектуальная истина станет предметом *непосредственного* восприятия (т.е. это еще не скоро).

3.1.4. Применение автоматизированных когнитивных технологий в качестве инструмента познания (на примере АСК-анализа и системы «Эйдос»)

3.1.4.1. АСК-анализ как автоматизированный метод научного познания

3.1.4.1.1. Кратко об АСК-анализе

Автоматизированный системно-когнитивный анализ (АСК-анализ) — это новый универсальный метод искусственного интеллекта, представляющий собой единственный в настоящее время вариант автоматизированного системного анализа, а именно, системный анализ, структурированный по базовым когнитивным операциям.

Известно, что системный анализ является одним из общепризнанных в науке методов решения проблем и многими учеными рассматривается вообще как методология научного познания. Однако как впервые заметил еще в 1984 году проф. И. П. Стабин³⁰ практическое применение системного анализа наталкивается на **проблему**, суть которой состоит в том, что методология системного анализа успешно применяется в сравнительно простых случаях, в которых в принципе можно обойтись и без нее, тогда как в реальных сложных ситуациях, она чрезвычайно востребован и у нее нет альтернатив, сделать это удается очень редко. Проф. И. П. Стабин первым предложил и путь решения этой проблемы, состоящий в автоматизации системного анализа, он же ввел и термин: «Автоматизированный системный анализ» (АСА).

В 2002 году в работе [6] Е.В.Луценко предложил структурировать системный анализ по базовым когнитивным операциям и назвал такой вариант автоматизированного системного анализа: «Автоматизированный системно-когнитивный анализ (АСК-анализ)». На тот момент по запросу: «Автоматизированный системно-когнитивный анализ» не выдавалось ни одного сайта, теперь же Яндекс находит 3 миллиона сайтов по этому запросу. Этих операций оказалось не так много и и их оказалось возможным математически описать с помощью системной теории информации [6]. Затем Е.В.Луценко разработал методику численных расчетов, включающую структуры данных и алгоритмы их обработки, а также программную реализацию: интеллектуальную систему «Эйдос», в настоящее время являющуюся программным инструментарием АСК-анализа [1-25].

³⁰ Стабин И.П., Моисеева В.С. Автоматизированный системный анализ.- М.: Машиностроение, 1984. –309 с.

3.1.4.1.2. Движение познания от эмпирических данных к информации, а от нее к знаниям

Из изложенного выше в разделе 2.2 можно сделать обоснованный вывод о том, что *АСК-анализ и система «Эйдос» обеспечивают движение познания от эмпирических данных к информации, а от нее к знаниям. По сути это движение от феноменологических моделей, описывающих явления внешне, к содержательным теоретическим моделям [3].*

3.1.4.1.3. Когнитивные функции

Когнитивные функции представляют собой отображение наглядном графическом виде количества информации, которое содержится различных значениях аргумента (T.e. значениях описательных шкал, значениях факторов) о различных значениях функции (градаций классификационных шкал, или классов) [7, 24-34]. Поэтому когнитивные функции отражают не только внешний вид функциональной зависимости, как в феноменологических моделях, но и смысл этой визуализации зависимости, т.е. являются средством причинноследственных зависимостей.

В работе [36] приведено теоретическое обоснование применения системно-когнитивных моделей вместо содержательных аналитических моделей сложных технических систем. Презентация к данной работе находится по адресу: http://ej.kubagro.ru/2016/07/upload/02.zip. В работе [37] приведен развернутый детальный пример такого применения, в т.ч. с использованием аппарата когнитивных функций.

3.1.4.1.4. Автоматизированный SWOT- и PEST-анализ

SWOT-анализ является широко известным и общепризнанным метод стратегического планирования. Однако это не мешает тому, что он подвергается критике, часто вполне справедливой, обоснованной и хорошо аргументированной. В результате критического рассмотрения SWOTанализа выявлено довольно много его слабых сторон (недостатков), источником которых является необходимость привлечения экспертов, в частности для оценки силы и направления влияния факторов. Ясно, что эксперты это делают неформализуемым путем (интуитивно), на основе своего профессионального опыта и компетенции. Но возможности экспертов имеют свои ограничения и часто по различным причинам они не могут и не хотят это сделать. Таким образом, возникает проблема проведения SWOT-анализа без привлечения экспертов. Эта проблема может решаться путем автоматизации функций экспертов, т.е. путем измерения силы и направления влияния факторов непосредственно на основе эмпирических данных. Подобная технология разработана давно, ей уже около 30 лет, но она малоизвестна – это интеллектуальная система «Эйдос». В статье на реальном численном примере подробно описывается

возможность проведения количественного автоматизированного SWOTанализа средствами АСК-анализа и интеллектуальной системы «Эйдос-X++» без использования экспертных оценок непосредственно на основе эмпирических данных. Предложено решение прямой и обратной задач SWOT-анализа [38]. PEST-анализ рассматривается как SWOT-анализ, с более детализированной классификацией внешних факторов [38].

Выводы, полученные с помощью Автоматизированного SWOT- и PEST-анализа можно непосредственно использовать для достижения целей, т.е. для управления. Это и означает, что АСК-анализ и его программный инструментарий — интеллектуальная система «Эйдос» обеспечивают создание непосредственно на основе эмпирических данных содержательных феноменологических моделей и их применение для решения задач идентификации, поддержки управляющих решений и исследования моделируемой предметной области.

3.1.4.1.5. Системно-когнитивные модели как содержательные эмпирические модели (выводы)

Из вышеизложенного можно обоснованно сделать следующие выводы:

- 1. Системно-когнитивные модели, создаваемые в АСК-анализе с помощью интеллектуальной системы «Эйдос» непосредственно на основе эмпирических данных представляют собой новый, ранее неизвестный класс моделей: содержательные феноменологические модели, которые промежуточное эмпирическими занимают положение между моделями и содержательными теоретическим феноменологическими обеспечивает визуализацию «Эйдос» Система моделируемом объекте, содержащихся в системно-когнитивных моделях, в форме развитой когнитивной графики, в частности в форме когнитивных функций, разнообразных когнитивных диаграмм и SWOT-диаграмм [38]. Необходимо особо подчеркнуть, что все эти диаграммы формируются не основе экспертных оценок, получаемых неформализуемым способом на основе интуиции, опыта и профессиональной компетенции, а на основе системно-когнитивных моделей, формируемых непосредственно на основе эмпирических данных.
- 2. АСК-анализ и система «Эйдос» обеспечивают движение познания от эмпирических данных к информации, а от нее к знаниям. По сути это движение от феноменологических моделей, описывающих явления внешне, к содержательным теоретическим моделям. Конечно, до создания теоретических моделей этот процесс не доходит, но он уменьшает разрыв или пропасть, существующую между эмпирическими феноменологическими моделями и содержательными теоретическим моделями. А значит и облегчает преодоление этого разрыва, т.е. подготавливает более благоприятную почву для разработки теоретических

моделей уже не на основе эмпирических моделей, а на основе системнокогнитивных моделей [36, 37]. Таким образом не смотря на то, что содержательными системно-когнитивные модели являются феноменологическими моделями и отражают смысловые причинноследственные связи между событиями реальной области они также содержательной теоретической интерпретации, требуют разработка специалиста-эксперта которой является содержательно делом разбирающегося в моделируемой предметной области.

3. Вместе с тем, опыт применения АСК-анализа и системы «Эйдос» показывает, что в ряде практически значимых случаев, таких, например, как принятие решений по рациональному выбору конструктивных особенностей и режимов работы сложных технических систем [36, 37], оказывается, что системно-когнитивных моделей в принципе достаточно для решения поставленных задач и в разработке содержательных теоретических моделей нет особой необходимости. А ранее, когда в исследователей разработчиков распоряжении были И ЛИШЬ феноменологические модели, а ранее такая необходимость была, т.к. эти модели не позволяли решать те задачи, которые решались лишь с применяем теоретических моделей.

3.1.4.2. Движение познания от частных и менее адекватных моделей объекта познания к более общим и более адекватным: принцип соответствия

3.1.4.2.1. Множественность адекватных моделей

Пусть у нас есть таблица с координатами точек: (X, Y), отражающая на эмпирическом уровне некоторую предметную область (результаты наблюдений или эксперимента). Эти точки образуют некое облако точек на плоскости. Спрашивается, как построить аналитическую модель этих эмпирических данных? Один из вариантов ответа на этот вопрос, реализованный в регрессионном анализе, состоит в том, чтобы провести некую кривую (тренд) таким образом, чтобы сумма квадрата отклонений от кривой до этих точек была минимальна? Другой вариант ответа на этот вопрос дает АСК-анализ [31, 39, 40]. Возможны и другие варианты. В реализации регрессионного анализа MS Excel пользователь сам выбирает функцию для аппроксимации эмпирических данных и нескольких Качество аппроксимации, T.e. достоверность количественно отражается в значении коэффициента детерминации \mathbb{R}^2 . При этом опыт показывает, что значения R^2 для разных функций иногда оказываются очень близкими, практически одинаковыми.

Это означает, что обычно на основе одних и тех же эмпирических данных возможно построить **несколько различных моделей**, которые отражают эти данные с практически одинаковой достоверностью. Поэтому возникает проблема выбора «единственной правильной модели»

из нескольких адекватных. Но если известна лишь единственная модель, то обычно она и принимается за «истинную». Понятно, что это весьма легкомысленно и вообще несерьезно, тем ни менее так чаще всего и происходит.

Для ученых, профессионально занимающихся разработкой моделей, все это вполне очевидно. Профессиональные разработчики моделей рассматривают множество различных вариантов повышения степени адекватности моделей [42] и не склонны какую-либо одну из них считать единственной истинной. Однако практически все остальные, в т.ч. ученые не разрабатывающие новых моделей, имеют такую склонность.

Совершенно аналогично, люди в различных состояниях и формах сознания строят различные модели реальности, включая и представления о пространстве, времени и движении, на основе эмпирических данных, поступающих от органов восприятия тех тел, которые поддерживает данные формы сознания [1, 2, 3]. В работе [3] автор обосновывает мысль о том, что чем выше адекватность и область действия модели реальности, созданной при некоторой определенной форме сознания, тем выше эта форма сознания, или, другими словами, чем выше форма сознания, тем более адекватная модель реальности создается при этой форме сознания. При развитии человеческого общества и повышении уровня наиболее массовой формы сознания повышается и степень адекватности моделей реальности, используемых человечеством [3, 41], и, соответственно, сменяют друг друга все более совершенные научные и мировоззренческие парадигмы, имеющие все более широкую область адекватности и лействия.

3.1.4.2.2. Принцип соответствия

Соотношение между новой более общей теорией и старой, описывающей какой-то частный случай новой теории, раскрывается методологическим принципом соответствия: новые теории включают адекватные аспекты старых теорий в виде своей части, но сохраняют адекватность в более широкой предметной области, чем старые модели, где старые модели теряют адекватность, в области действия, в которой старая теория была адекватна, новая теория дает практически те же самые предсказания, что и старая теория.

3.1.4.3. Движение познания от моделей низкого уровня формализации к моделям более высокого уровня формализации

Существуют различные типы *измерительных шкал* (номинальные, порядковые и числовые) [35] и различные виды *моделей*.

В процессе познания степень формализации моделей постоянно возрастает. Зачем нужно повышать степень формализации? Прежде всего

для того, чтобы передать знания другим и использовать их для опредмечивания, применения в технологиях и обществе. Мы ведь не можем передать знания в интуитивной форме. Этапы формализации: вербализация и т.д. мысль произнесенная есть ложь. Необходимо отметить, что на каждом этапе повышения уровня формализации происходит необратимая потеря информации.

3.1.5. Краткий обзор научных работ автора по когнитивной ветеринарии

В монографии подробно рассмотрено решение ряда задач когнитивной ветеринарии, которые аннотируются ниже.

1. Когнитивная ветеринария - ветеринария цифрового общества: дефиниция базовых понятий.

Раздел 3.1 основан на работе [1]. Ссылки на литературные источники в данном разделе даны по списку литературы работы [1].

Есть много мнений по вопросу о том, в каком обществе мы живем в начале XXI века. Ранее считали, что это постиндустриальное общество. Было даже такое мнение, что это будет общество развитого социализма или даже коммунистическое общество. Потом мнение по этому поводу Сначала современное общество изменилось. стали называть информационным обществом, а затем обществом, основанным на знаниях. Последняя новость в этой области состоит в том, что, по-видимому, современное общество – это цифровое общество, т.е. общество, цифровых технологиях, цифровых коммуникациях, основанное на цифровых технологиях обработки и передачи информации, и цифровых технологиях искусственного интеллекта. В цифровом обществе и науки перейти цифровым интеллектуальным должны технологиям исследования. В частности возникает вопрос о том, не должна ли и ветеринария в цифровом обществе стать когнитивной ветеринарией. Развернутому и аргументированному (по мнению авторов) ответу на этот вопрос и посвящена данная работа. Методология и терминология в данной новой области еще не устоялась и не является общепринятой. Поэтому в данной работе много внимания уделено логике и методологии научного познания, терминологическим вопросам и дефинициям понятий.

2. Формирование семантического ядра ветеринарии путем АСКанализа паспортов научных специальностей ВАК РФ и автоматическая классификация текстов по направлениям науки.

Раздел 3.2 основан на работе [2]. Ссылки на литературные источники в данном разделе даны по списку литературы работы [2].

Данная работа является продолжением серии работ автора по когнитивной ветеринарии. Настоящее время характеризуется появлением в открытом доступе огромных объемов текстов на различных языках, сгенерированных людьми. В настоящее время эти тексты накапливаются в

различных электронных библиотеках и библиографических базах данных (WoS, Скопус, РИНЦ и др), а также просто в Internet на различных сайтах. Все эти тексты имеют конкретных авторов, датировку и могут относиться одновременно ко многим не альтернативным категориям и жанрам, в частности: учебные; научные; художественные; политические; новостные; чаты; форумы и многие другие. Большой научный и практический интерес представляет решение обобщенной задачи атрибуции текстов, т.е. такого исследования этих текстов, при котором определялись бы их вероятные принадлежность текстов датировка создания, авторы, ЭТИХ перечисленным выше обобщенным группам или жанрам, а также оценка сходства- различия авторов и текстов по их содержанию, выделение в текстах ключевых слов и т.п. и т.д. Для решения всех этих задач необходимо сформировать обобщенные лингвистические образы текстов по группам (классам), т.е. сформировать семантические ядра классов. Частным случаем этой задачи является создание семантических ядер по научным специальностям ВАК РФ различным автоматическая классификация научных текстов по направлениям науки. Традиционно эта задача решается диссертационными советами, т.е. экспертами, на основе экспертных оценок, т.е. неформализованным путем, на основе опыта, интуиции и профессиональной компетенции. Однако традиционный подход имеет ряд довольно серьезных недостатков, накладывающих на качество и объемы анализа существенные ограничения. В настоящее время уже есть все основания рассматривать эти ограничения как неприемлемые, т.к. их вполне можно преодолеть. Таким образом, налицо проблема, пути решения которой и являются предметом рассмотрения в данной статье. актуальными является исследователей Следовательно, усилия разработчиков по их преодолению. Поэтому целью работы является разработка автоматизированной технологии (метода и инструментария), а также методики их применения для формирования семантического ядра ветеринарии путем автоматизированного системно-когнитивного анализа научных специальностей BAK РΦ паспортов автоматической классификация текстов по направлениям науки. Приводится развернутый численный пример решения поставленной проблемы на реальных данных.

3. Синтез семантических ядер научных специальностей ВАК РФ и автоматическая классификации статей по научным специальностям с применением АСК-анализа и интеллектуальной системы "Эйдос" (на примере Научного журнала КубГАУ и его научных специальностей: механизации, агрономии и ветеринарии).

Раздел 3.3 основан на работе [3]. Ссылки на литературные источники в данном разделе даны по списку литературы работы [3].

14 января 2019 года на сайте BAK PФ http://vak.ed.gov.ru/87 появилась информация: «Об уточнении научных специальностей и соответствующих им отраслей науки, по которым издания входят в

Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук». Сообщается, что согласно рекомендации ВАК для остальных изданий, входящих в Перечень по группам научных специальностей, работа по уточнению научных специальностей и отраслей науки будет продолжена в 2019 году. Данная работа является продолжением серии автора по когнитивной лингвистике. В ней предлагается инновационная интеллектуальная технология для автоматизации решения применением сформулированной ВАК РΦ выше. автоматизированного системно-когнитивного анализа (АСК-анализ) и его программного инструментария – интеллектуальной системы «Эйдос» непосредственно на основе официальных текстов паспортов научных специальностей ВАК РФ созданы их семантические ядра, а затем реализована автоматическая классификация научных текстов (статей, монографий, учебных пособий и т.д.) по специальностям и группам Традиционно специальностей ВАК РΦ. решается эта задача диссертационными советами, а также редакционными советами научных изданий, экспертами, на основе экспертных оценок, неформализованным путем, основе интуиции на опыта, профессиональной компетенции. Однако, традиционный подход имеет ряд довольно серьезных недостатков, накладывающих на качество и объемы анализа существенные ограничения. Следовательно, актуальными является разработчиков исследователей И ПО преодолению ограничений. В настоящее время уже есть все основания рассматривать эти ограничения как неприемлемые, т.к. их не только нужно, но и вполне возможно преодолеть. Таким образом, налицо проблема, решение которой и являются предметом рассмотрения в данной статье. Приводится развернутый численный пример решения поставленной проблемы на реальных данных.

4. АСК-анализ в ветеринарии (на примере разработки диагностических тестов).

Раздел 3.4 основан на работе [4]. Ссылки на литературные источники в данном разделе даны по списку литературы работы [4].

В статье рассмотрено применение Автоматизированного системнокогнитивного анализа (АСК-анализ) и его программного инструментария интеллектуальной технологи «Эйдос» для реализации уже разработанных ветеринарных диагностических без И медицинских тестов программирования в форме, удобной для индивидуального и массового тестирования, анализа его результатов и выработки индивидуальных и групповых рекомендаций. Возможно объединение нескольких ветеринарных тестов в один супертест.

5. Реализация тестов и супертестов для ветеринарной и медицинской диагностики в среде системы искусственного интеллекта "Эйдос-X++" без программирования.

Раздел 3.5 основан на работе [5]. Ссылки на литературные источники в данном разделе даны по списку литературы работы [5].

В статье рассмотрено применение интеллектуальной технологи «Эйдос» для реализации уже разработанных ветеринарных и медицинских диагностических тестов без программирования в форме, удобной для индивидуального и массового тестирования, анализа его результатов и выработки индивидуальных и групповых рекомендаций. Возможно объединение нескольких тестов в супертест.

6. Агломеративная когнитивная кластеризация нозологических образов в ветеринарии.

Раздел 3.6 основан на работе [6]. Ссылки на литературные источники в данном разделе даны по списку литературы работы [6].

В статье на небольшом численном примере рассматривается сходство и различие нозологических образов в ветеринарии с применением агломеративной кластеризации, реализованного автоматизированном системно-когнитивном анализе (АСК-анализ). Этот метод получил название: «Агломеративная когнитивная кластеризация». Этот метод отличается от известных традиционных тем, что: а) в нем параметры обобщенного образа кластера вычисляются не как средние от исходных объектов (классов) или их центр тяжести, а определяются с помощью той же самой базовой когнитивной операции АСК-анализа, которая применяется и для формирования обобщенных образов классов на объектов и которая действительно примеров обеспечивает обобщение; б) в качестве критерия сходства используется не евклидово расстояние или его варианты, а интегральный критерий неметрической природы: «суммарное количество применение которого теоретически корректно и дает хорошие результаты в неортонормированных пространствах, которые обычно и встречаются на практике; в) кластерный анализ проводится не на основе исходных переменных, матриц частот или матрицы сходства (различий), зависящих от единиц измерения по осям, а в когнитивном пространстве, в котором по всем осям (описательным шкалам) используется одна единица измерения: количество информации, и поэтому результаты кластеризации не зависят от исходных единиц измерения признаков объектов. Все это позволяет получить результаты кластеризации, понятные специалистам поддающиеся содержательной интерпретации, хорошо согласующиеся с оценками экспертов, их опытом и интуитивными ожиданиями, что часто представляет собой проблему для классических методов кластеризации.

7. Агломеративная когнитивная кластеризация симптомов и синдромов в ветеринарии.

Раздел 3.7 основан на работе [7]. Ссылки на литературные источники в данном разделе даны по списку литературы работы [7].

В статье на небольшом численном примере рассматривается сходство и различие симптомов и синдромов по их диагностическому смыслу, т.е. по той информации, которую они содержат о принадлежности состояний животных к различным нозологическим образам. Эта задача решается для ветеринарии с применением нового метода агломеративной кластеризации, реализованного в автоматизированном системно-когнитивном анализе (АСК-анализ). Этот метод кластеризации отличается от известных традиционных тем, что: а) в нем параметры обобщенного образа кластера вычисляются не как средние от исходных объектов (симптомов) или их центр тяжести, а определяются с помощью той же самой базовой когнитивной операции АСК-анализа, которая применяется и для формирования обобщенных образов классов на основе примеров объектов и которая действительно корректно обеспечивает обобщение; б) в качестве критерия сходства используется не евклидово расстояние или его варианты, а интегральный критерий неметрической природы: «суммарное количество информации», применение которого хорошие теоретически корректно дает результаты И неортонормированных пространствах, которые как правило и встречаются на практике; в) кластерный анализ проводится не на основе исходных переменных, матриц частот или матрицы сходства (различий), зависящих от единиц измерения по осям (измерительным шкалам), а в когнитивном пространстве, в котором по всем осям используется одна единица измерения: количество информации, и поэтому результаты кластеризации не зависят от исходных единиц измерения признаков объектов. Все это позволяет получить результаты кластеризации, понятные специалистам и поддающиеся содержательной интерпретации, хорошо согласующиеся с оценками экспертов, их опытом и интуитивными ожиданиями, что часто представляет собой проблему для классических методов кластеризации.

8. АСК-анализ и классификация пород крупного рогатого скота.

Раздел 3.8 основан на работе [8]. Ссылки на литературные источники в данном разделе даны по списку литературы работы [8].

На сайте мясной академии http://meatinfo.ru по адресу: http://meatinfo.ru/info/show?id=197 размещена сравнительная таблица пород крупного рогатого скота по 8 показателям, из которых 2 текстовых и 6 числовых. У хозяйственников возникает естественный вопрос о том, какие из этих пород сходны по всей системе характеризующих их показателей, а какие различаются и в какой степени. Возникает также вопрос о том, какие показатели сходны и отличаются по смыслу и на сколько. Решению этих задач и посвящена данная статья. Результаты исследования могут быть

использованы всеми желающими, благодаря тому, что Универсальная автоматизированная система «Эйдос», являющаяся инструментарием Автоматизированного системно-когнитивного анализа, находится в полном открытом бесплатном доступе на сайте автора по адресу: http://lc.kubagro.ru/aidos/_Aidos-X.htm, а численный пример решения поставленных задач размещен как облачное Эйдос-приложение №131

9. АСК-анализ антибиотиков в ветеринарии.

Раздел 3.9 основан на работе [9]. Ссылки на литературные источники в данном разделе даны по списку литературы работы [9].

Антибактериальные химиотерапевтические препараты, к которым относятся антибиотики и синтетические противомикробные средства, широко применяются в ветеринарии для профилактики и лечения заболеваний, микроорганизмами. Антибактериальные вызываемых средства можно классифицировать по типу действия и химической структуре. Известно также, что при применении нескольких препаратов в сочетании друг с другом они взаимодействуют внутри организма друг с другом, что может приводить к усилению или ослаблению их действия. По этим причинам представляет научный и практический интерес разработка классификации антибиотиков по их характеристикам и принципу действия (задача 1), а также по взаимной совместимости (задача 2). Эти задачи решаются в статье с применением нового метода агломеративной кластеризации, реализованного в автоматизированном системно-когнитивном анализе (АСК-анализ). Этот метод кластеризации имеет ряд преимуществ перед известными традиционными методами кластеризации. Эти преимущества позволяют получить результаты кластеризации, понятные специалистам и поддающиеся содержательной интерпретации, хорошо согласующиеся с оценками экспертов, их опытом и интуитивными ожиданиями, что часто представляет собой проблему для классических методов кластеризации. В статье приводятся подробные численные примеры решения двух поставленных задач. Универсальная автоматизированная система «Эйдос», являющаяся инструментарием АСК-анализа, находится в полном открытом бесплатном доступе на сайте адресу: http://lc.kubagro.ru/aidos/_Aidos-X.htm. решения задач ветеринарии с применением примеры технологий искусственного интеллекта размещены как облачные Эйдос-приложения и доступны всем желающим.

10. АСК-анализ влияния пробиотиков в рационах на телосложение бычков.

Раздел 3.10 основан на работе [10]. Ссылки на литературные источники в данном разделе даны по списку литературы работы [10].

Статья посвящена применению автоматизированного системно-когнитивного анализа для исследования пробиотиков на рост бычков, причем в качестве индикаторов роста использована не живая масса, а

индексы, характеризующие форму телосложения и пропорции тела животных. Приводится подробный численный пример решения поставленной задачи на реальных данных.

11. Разработка ветеринарного теста для диагностики желудочнокишечных заболеваний лошади на основе данных репозитория UCI с применением ACK-анализа.

Раздел 3.11 основан на работе [11]. Ссылки на литературные источники в данном разделе даны по списку литературы работы [11].

В данной статье кратко рассматривается новый инновационный (доведенный до уровня, обеспечивающего практическое использование) искусственного интеллекта: автоматизированный когнитивный анализ (АСК-анализ) и его программный инструментарий интеллектуальная система «Эйдос». Приводится подробный численный пример решения, демонстрирующий технологию создания ветеринарного диагностического теста желудочно-кишечных заболеваний лошади. В качестве исходных данных использованы данные репозитория UCI. предоставленные Мэри Маклиш и Мэтт Сесиль (Отдел компьютерных наук Гуэлфский университет, Онтарио, Канада N1G 2W1, при поддержке спонсора: Уилла Тейлора. Разработанный тест использован для решения задач диагностики, поддержки принятия решений и исследования моделируемой предметной области путем исследования ее модели. Результаты исследования могут быть использованы всеми желающими, благодаря тому, что Универсальная автоматизированная система «Эйдос», инструментарием АСК-анализа, являющаяся находится полном бесплатном доступе на сайте автора адресу: http://lc.kubagro.ru/aidos/_Aidos-X.htm, а численные примеры решения задач ветеринарии с применением технологий искусственного интеллекта размещены как облачное Эйдос-приложение № 129

12. Математическое и численное моделирование взаимосвязи морфологического, биохимического и микроэлементного состава крови бычков герефордской породы и их размеров.

Раздел 3.12 основан на работе [12]. Ссылки на литературные источники в данном разделе даны по списку литературы работы [12].

Исследователями получены морфологическому, данные ПО биохимическому и микроэлементному составу крови бычков герефордской породы различных размеров. В этой связи у ученых и хозяйственников возникает три естественных вопроса: 1) возможно ли по этим показателям крови прогнозировать размеры и тем самым мясную продуктивность бычков; каковы сила и направление влияния тех или иных значений показателей крови на размер и вес бычков; какие показатели крови сходны по смыслу, а какие отличаются и на сколько (в какой степени). Аргументированным ответам на ЭТИ вопросы путем применения современных методов математического и численного моделирования для решения соответствующих задач и посвящена данная статья. Результаты исследования могут быть использованы всеми желающими, благодаря Универсальная автоматизированная система «Эйдос», TOMY, что инструментарием Автоматизированного системноявляющаяся когнитивного анализа, находится в полном открытом бесплатном доступе на сайте автора по адресу: http://lc.kubagro.ru/aidos/ Aidos-X.htm, а численный пример решения поставленных задач размещен как облачное Эйдос-приложение №133.

13. Когнитивная информационно-измерительная квалиметрическая система для определения содержания жира и белка в коровьем молоке по параметрам тензиограмм динамического поверхностного натяжения на границе раздела молоко/воздух.

Раздел 3.13 основан на работе [13]. Ссылки на литературные источники в данном разделе даны по списку литературы работы [13].

Рядом авторов (Милаёва И.В., Зайцев С.Ю., Довженко Н.А., Царьков Д.В., Царькова М.С., 2015) предложена регрессионная модель и способ косвенного измерения содержания жира и белка в коровьем молоке по его динамическому поверхностному натяжению, имеющий ряд преимуществ перед традиционным подходом. Эта модель отражает объективно существующие взаимосвязи между содержанием жира и белка в коровьем молоке и параметрами тензиограмм динамического поверхностного натяжения на границе раздела молоко/воздух. Эти взаимосвязи выявлены авторами способа путем математической обработки 112 эмпирических проб. Для этого ими был применен регрессионный и корреляционный анализ (в MS Excel). Так как содержание жира и белков в молоке во многом определяет его качество, то данная задача относится к задачам квалиметрии. Однако, данная задача квалиметрии относится также к распознавания образов (многопараметрической типичным задачам типизации и системной идентификации), и, поэтому вполне возможно, что качество ее решения может быть повышено путем применения методов искусственного интеллекта, в частности когнитивных и информационных технологий. Тем более, что эти взаимосвязи имеют довольно сложный характер. Применение интеллектуальных технологий открывает также дополнительные возможности исследования моделируемой ДЛЯ предметной области путем исследования ее модели. Все это представляет большой научный и практический интерес как для ученых исследователей, а так и для практиков. Для решения всех этих задач в работе применен Автоматизированный системно-когнитивный анализ (АСК-анализ) и его программный инструментарий – интеллектуальная система «Эйдос». Подробно рассмотрен численный пример, основанный на реальных данных.

3.1.6. Заключение

Необходимо отметить, что системно-когнитивные модели, разработанные в системе «Эйдос», могут быть применены для решения практических задач с применением той же системы «Эйдос», в которой они созданы, причем это применение возможно в адаптивном режиме, т.е. их можно совершенствовать в процессе эксплуатации, адаптировать к изменениям предметной области, локализовать или районировать для других регионов. Эти уникальные возможности обеспечиваются тем, что представляет система «Эйдос» собой не только среду интеллектуальных эксплуатации приложений, но является инструментом их создания и адаптации.

Таким образом АСК-анализ и система «Эйдос» представляют собой новый инновационный, т.е. реально доведенный до возможности практического применения, метод искусственного интеллекта, который обоснованно может рассматриваться как универсальный инструмент решения всех тех задач в области ветеринарии (и других наук), для решения которых используется естественный интеллект. Причем это инструмент, многократно увеличивающий возможности естественного интеллекта, примерно также, как микроскоп и телескоп многократно увеличивает возможности естественного зрения, естественно только в том случае, если оно есть. Поэтому, конечно, этих задач огромное количество.

В качестве перспектив можно было бы отметить в частности решение следующих задач ветеринарии с применением автоматизированного системно-когнитивного анализа:

- поддержка принятия решений по выбору антибактериальных препаратов в зависимости от характера микробной флоры;
- поддержка принятия решений по определению дозы и пути введения препаратов группы пенициллина;
- поддержка принятия решений по определению дозы и пути введения цефалоспоринов;
- поддержка принятия решений по выбору антибактериальных препаратов с учетом основных токсических и аллергических реакций на антибактериальные препараты;
- исследование взаимодействия антибактериальных препаратов с другими препаратами при приеме внутрь и поддержка принятия решений по выбору антибактериальных препаратов с учетом результатов этих исследований.

Область ветеринарии, в которой перечисленные выше и другие задачи решаются с применением системно-когнитивного анализа, программным инструментарием которого в настоящее время является система «Эйдос», предлагается назвать «Когнитивной ветеринарией» или шире «Математической ветеринарией», по аналогии с когнитивной

экономикой, математической экономикой (08.00.13), Прикладной и математической лингвистикой (10.02.21), когнитивной лингвистикой и т.д.

Надо отметить, что все сказанное в данной статье практически без изменений относится не только к ветеринарии, но и к любой науке, которая еще недостаточно пользуется математикой и интеллектуальными технологиями. Ясно, что науки будут развиваться в этом направлении и со временем появятся основания назвать их, например когнитивной агрономией, когнитивной ампелографией, когнитивным овощеводством, когнитивным плодоводством, когнитивной механизацией и т.д.

Эта идея находится в русле Указа Президента РФ от 7 июля 2011 г. N 899 "Об утверждении приоритетных направлений развития науки, технологий и техники в Российской Федерации и перечня критических технологий Российской Федерации", в котором под п.8 указаны Нано-, био-, информационные, когнитивные технологии 31.

Этим и другим применениям способствует и то, что система «Эйдос» является мультиязычной интеллектуальной on-line средой для обучения и научных исследований [2, 3]³² и находится в полном открытом бесплатном доступе (причем с подробно комментированными актуальными исходными текстами: http://lc.kubagro.ru/_AIDOS-X.txt) на сайте автора по адресу: http://lc.kubagro.ru/aidos/_Aidos-X.htm. [2-11].

Численные примеры решения задач ветеринарии с применением технологий искусственного интеллекта размещены как облачные Эйдосприложения под номерами: 100, 125, 126, 127, 128, 131, 133, 134, 144 и доступны всем желающим в режиме 1.3 системы «Эйдос».

Конечно, представленный в статье уровень исследования относится хотя и к развитому, но эмпирическому уровню, т.е. это просто наблюдаемые факты, эмпирические закономерности и в лучшем случае, условии подтверждения полученных результатов исследователями, может подняться до уровня эмпирического закона. Для перехода на теоретический уровень познания необходимо выдвинуть гипотезы содержательной интерпретации полученных результатов (которые может выдвинуть только специалист в области ветеринарии), объясняющие внутренние механизмы наблюдаемых закономерностей. Потом необходимо подтвердить, что эти научные гипотезы имеют прогностическую силу, T.e. позволяют обнаружить неизвестные явления, и тогда эти гипотезы переходят в статус научной теории. Эта теория позволяют обобщить эмпирический закон до уровня научного закона [1-5].

http://lc.kubagro.ru/aidos/Presentation_Aidos-online.pdf

 $^{^{31}}$ Отметим, что все приведенные выше аргументы введения научного понятия: «когнитивная ветеринария» применимы и к другим направлениям науки, например: «когнитивная агрономия», «когнитивная экономика» и т.д.. Автор пытался развивать когнитивную математику [10] и когнитивную теорию управления [10], а также применять их в других областях науки и практики.

3.2. Формирование семантического ядра ветеринарии путем АСК-анализа паспортов научных специальностей ВАК РФ и автоматическая классификация текстов по направлениям науки

Данная работа является продолжением серии работ автора по когнитивной ветеринарии. Настоящее время характеризуется появлением в открытом доступе огромных объемов текстов на различных языках, сгенерированных людьми. В настоящее время эти тексты накапливаются в различных электронных библиотеках и библиографических базах данных (WoS, Скопус, РИНЦ и др), а также просто в Internet на различных сайтах. Все эти тексты имеют конкретных авторов, датировку и могут относиться одновременно ко многим не альтернативным категориям и частности: учебные; научные; художественные; политические; новостные; чаты; форумы и многие другие. Большой научный и практический интерес представляет решение обобщенной задачи атрибуции текстов, т.е. такого исследования этих текстов, при котором определялись бы их вероятные авторы, датировка создания, принадлежность этих текстов к перечисленным выше обобщенным группам или жанрам, а также оценка сходства- различия авторов и текстов по их содержанию, выделение в текстах ключевых слов и т.п. и т.д. Для решения всех этих задач необходимо сформировать обобщенные лингвистические образы текстов по группам (классам), т.е. сформировать семантические ядра классов. Частным случаем этой задачи является создание семантических ядер по различным научным специальностям BAK $P\Phi$ и автоматическая классификация научных текстов по направлениям науки. Традиционно эта задача решается диссертационными советами, т.е. экспертами, на основе экспертных оценок, т.е. неформализованным путем, на основе опыта, интуиции и профессиональной компетенции. Однако традиционный подход имеет ряд довольно серьезных недостатков, накладывающих на качество и объемы анализа существенные ограничения. В настоящее время уже есть все основания рассматривать эти ограничения как неприемлемые, т.к. их вполне можно преодолеть. Таким образом, налицо проблема, пути решения которой и являются предметом рассмотрения в данной статье. Следовательно, актуальными является усилия исследователей и разработчиков по их преодолению. Поэтому целью работы является разработка автоматизированной технологии (метода и инструментария), а также методики их формирования семантического для ядра ветеринарии автоматизированного системно-когнитивного анализа паспортов научных специальностей ВАК РФ и автоматической классификация текстов по направлениям науки. Приводится развернутый численный пример решения поставленной проблемы на реальных данных.

Раздел 3.2 основан на работе [2]. Ссылки на литературные источники в данном разделе даны по списку литературы работы [2].

3.2.1. Введение

Данная работа является продолжением серии работ автора по когнитивной ветеринарии [26-33] и когнитивному анализу текстов [1-20]. Настоящее время характеризуется появлением в открытом доступе огромных объемов текстов на различных языках, сгенерированных

людьми. В настоящее время эти тексты накапливаются в различных электронных библиотеках и библиографических базах данных (WoS, Сопус, РИНЦ и др), а также просто в Internet на различных сайтах.

Все эти тексты имеют конкретных авторов, датировку и могут относится одновременно ко многим не альтернативным категориям и жанрам, в частности: учебные; научные; художественные; политические; новостные; чаты; форумы и многие другие.

Большой научный и практический интерес представляет решение обобщенной задачи атрибуции текстов, т.е. такого исследования этих текстов, при котором определялись бы их вероятные авторы, датировка создания, принадлежность этих текстов к перечисленным выше обобщенным группам или жанрам, а также оценка сходства- различия авторов и текстов по их содержанию, выделение в текстах ключевых слов и т.п. и т.д. [1-5].

Для решения всех этих задач необходимо сформировать обобщенные лингвистические образы текстов по группам (классам), т.е. сформировать семантические ядра классов. Важным частным случаем этой задачи является создание семантических ядер по различным научным специальностям ВАК РФ и автоматическая классификация научных текстов по направлениям науки. При приеме новой диссертации к защите в диссертационном совете обязательно проводится голосование членов совета по вопросу о том, соответствует ли данная диссертация направлению науки, по которому она представлена к защите и профилю диссертационного совета. При этом члены диссертационного совета решают ту же самую задачу.

Традиционно эта задача решается диссертационными советами, т.е. экспертами, на основе экспертных оценок, т.е. неформализованным путем, на основе опыта, интуиции и профессиональной компетенции. Однако традиционный подход имеет ряд довольно серьезных недостатков, накладывающих на качество и объемы анализа существенные ограничения. В настоящее время уже есть все основания рассматривать эти ограничения как неприемлемые, т.к. их вполне можно преодолеть. Таким образом налицо **проблема**, пути решения которой и являются предметом рассмотрения в данной статье.

Какие же конкретно основные недостатки традиционного подхода?

<u>Во-первых</u>, количество экспертов-аналитиков по контент-анализу и атрибуции текстов довольно ограничено.

<u>Во-вторых</u>, время экспертов стоит дорого, а производительность их работы не очень высока. Поэтому задача атрибуции текстов ими решается неточно и в ограниченном объеме.

<u>В-третьих</u>, обычно эксперты это очень уважаемые и довольно пожилые люди, имеющие ограниченную производительность и не очень ладящие с компьютерами. Поэтому обычно эксперты лично не работают с

компьютерами и им для поддержки этой работы нужен специалист, обычно когнитолог (инженер по знаниям), который извлекает знания эксперта и повышает степень их формализации до уровня, который позволяет внести их в интеллектуальные системы.

<u>В-четвертых</u>, часто эксперт не способен формализовать свои знания и оценки даже до уровня вербализации, т.е. сформулировать их словами. Иначе говоря, все оценки производятся в качественных порядковых шкалах с небольшим числом градаций. Никакой речи об использования количественных шкал с большим числом интервальных числовых значений не идет.

В-пятых, эксперты часто не хотят искренне сообщать свои знания и методы принятия решений, т.к. иногда в особо щепетильных случаях такие откровения экспертов можно было бы квалифицировать как чистосердечное признание, снижающее тяжесть наказания. Понятно, что в подобных случаях эксперт вместо реальных используемых на практике знаний и подходов сообщит когнитологу вполне ожидаемые и приемлемые с точки зрения морально-этических норм, а также гражданского и уголовного законодательства.

Мы видим, что недостатки традиционного решения являются существенными, усилия довольно значит, исследователей a разработчиков по их преодолению являются актуальными. Поэтому целью работы является разработка автоматизированной технологии (метода и инструментария), а также методики их применения для ветеринарии формирования семантического ядра автоматизированного системно-когнитивного анализа паспортов научных специальностей ВАК РФ и автоматической классификация текстов по направлениям науки. В работе приводится развернутый численный пример решения поставленной проблемы на реальных данных.

Для достижения поставленной цели необходимо решить следующие **задачи**, которые получаются путем декомпозиции цели:

Задача 1: сформулировать идею и концепцию решения проблемы;

<u>Задача 2:</u> обосновать выбор метода и инструмента решения проблемы;

<u>Задача 3:</u> применить выбранный метод и инструмент для решения поставленной проблемы, т.е. выполнить следующие этапы:

- 3.1. Когнитивная структуризация предметной области.
- 3.2. Формализация предметной области.
- 3.3. Синтез и верификация модели.
- 3.4. Повышение качества модели и выбор наиболее достоверной модели.
- 3.5. Решение в наиболее достоверной модели задач диагностики (классификации, распознавания, идентификации), поддержки принятия

решений и исследования моделируемой предметной области путем исследования ее модели.

Задача 4: описать эффективность предложенного решения проблемы.

<u>Задача 5:</u> рассмотреть ограничения и недостатки предложенного решения проблемы и перспективы его развития путем их преодоления этих ограничений и недостатков.

Ниже кратко рассмотрим решение этих задач.

3.2.2. Материалы и методы

3.2.2.1. Идея и концепция решения проблемы (задача 1)

Идея решения проблемы состоит в применении для этой цели современных IT-технологий, особенно интеллектуальных технологий, которых просто не существовало в период разработки традиционного подхода.

Концепция решения проблемы конкретизирует сформулированную выше идею и заключается в применении технологий искусственного интеллекта для создания приложений, обеспечивающих глубокую атрибуцию текстов (термин автора).

Суть предлагаемой концепции состоит в применении теории информации для того, чтобы рассчитать какое количество информации содержится в том или ином слове или его лемме о принадлежности текста с этим словом (леммой) к определенному классу.

Текст считается тем более релевантным классу, чем больше суммарное количество информации, содержащееся в словах (леммах) этого текста о принадлежности к этому классу.

Классы можно сравнивать друг с другом по тому, какое количество информации о принадлежности к ним содержат слова (леммы), встречающиеся в текстах этих классов.

3.2.2.2. Обоснование выбора метода и инструмента решения проблемы (задача 2)

Кратко метод АСК-анализа, суть математической модели АСК-анализа, включая частные и интегральные критерии, интеллектуальная система «Эйдос», являющаяся единственным в настоящее время программным инструментарием АСК-анализа, описаны в разделе 2 данной работы, поэтому здесь не приводятся.

3.2.2.3. Применение системы «Эйдос» для создания интеллектуального приложения (задача 3)

Решение задачи 3 предполагает выполнение следующих этапов, стандартных для АСК-анализа:

- когнитивная структуризация предметной области;
- формализация предметной области;

- синтез и верификация модели;
- повышение качества модели и выбор наиболее достоверной модели
- решение в наиболее достоверной модели задач диагностики (классификации, распознавания, идентификации), поддержки принятия решений и исследования моделируемой предметной области путем исследования ее модели.

Все эти этапы АСК-анализа, за исключением 1-го, автоматизированы в системе «Эйдос». Рассмотрим их в порядке исполнения.

2.2.2.3.1. Когнитивно-целевая структуризация предметной области

На этапе когнитивно-целевой структуризации предметной области мы решаем на качественном уровне, что будем рассматривать в качестве факторов, действующих на моделируемый объект (причин), а что в качестве результатов действия этих факторов (последствий).

В данной работе мы собираемся на основе паспортов различных специальностей РΦ, сформировать обобщенные научных ВАК т.е. специальностей, лингвистические образы ЭТИХ сформировать семантические ядра классов по различным направлениям науки, соответствующих этим специальностям.

При этом нас различные специальности интересуют в различной степени (это связано просто с местом работы автора), поэтому мы специальности, связанные с информационными технологиями, механизацией, агрономией и ветеринарией рассмотрим более подробно (полные шифры специальностей), а остальные на более обобщенном уровне (первые две пары цифр шифров специальностей).

Соответственно, в качестве классификационных шкал мы используем шифры и наименования научных специальностей ВАК РФ, а в качестве описательных шкал слова (точнее их леммы) этих текстов (мемы на в данном исследовании не используются, хотя применяемая технология это позволяет сделать).

2.2.2.3.2. Формализация предметной области и описание исходных данных

Исходные данные представляют собой текстовые файлы паспортов научных специальностей ВАК РФ, взятые непосредственно с сайта ВАК на странице: http://vak.ed.gov.ru/316. Эти файлы скачаны с сайта ВАК вручную по приведенным на нем прямым ссылкам.

Характеристика файлов исходных данных. Паспорта специальностей (их 429) представляют собой doc-файлы MS Word (рисунок 3a)..

```
Fig. 1.1

| Control | Cont
```

Рисунок За. Файлы исходных данных

Имена этих файлов обычно, но не всегда, состоят из шифра специальности и ее наименования, причем наименования специальностей часто являются сокращенными, а иногда и вообще отсутствуют (есть только шифры). Одна специальность (09.00.04 Эстетика) повторяется на сайте ВАК два раза.

Для системы «Эйдос» в настоящее время исходные текстовые файлы должны быть в кодировке DOS OEM 866 (ASCII). Поэтому исходные вордовские doc-файлы, скачаннеы с сайта BAK, были сначала преобразованы в DOS-кодировку Windows (ANSI) (с помощью Total Doc Converter 5.1.0.191), а затем перекодированы из ANSI в ASCII (с помощью DS text converter).

Затем наименования скачанных файлов паспортов специальностей были приведены в полное соответствие с теми, которые приведены на сайте ВАК. Это было сделано не вручную, а с помощью специально для этого разработанной автором небольшой программы, исходный текст которой приводится ниже:

```
FUNCTION Main()

CLOSE ALL

USE PassportsNameVAK EXCLUSIVE NEW

aPassportsNameVAK := {}

DBGOTOP()

DO WHILE .NOT. EOF()

mPassportsNameVAK = ALLTRIM(FIELDGET(2))

mPassportsNameVAK = STRTRAN (mPassportsNameVAK, ',', ';')

mPassportsNameVAK = SUBSTR (mPassportsNameVAK, 1, 6)+'00, '+mPassportsNameVAK+'.txt'

AADD (aPassportsNameVAK, mPassportsNameVAK)

DBSKIP(1)

ENDDO

mCountTxt = ADIR("*.TXT") // Кол-во исходных ТХТ-файлов с паспортами специальностей

PRIVATE aFN[mCountTxt]

ADIR("*.txt", aFN)
```

```
aFileName := {}
FOR j=1 TO mCountTxt
    IF LEN(ALLTRIM(SUBSTR(aFN[j], 9, 1))) = 0
       AADD(aFileName, ALLTRIM(aFN[j]))
NEXT
*LB_Warning(aFileName)
mCountTxt = LEN(aFileName)
FOR j=1 TO mCountTxt
    mVAKspecialtyCipher = SUBSTR(aFileName[j], 1, 8)
    FOR i=1 TO mCountTxt
        mPos = AT(mVAKspecialtyCipher, aPassportsNameVAK[i])
        IF mPos > 0
           mOldName = aFileName[j]
           mNewName = ConvToAnsiCP(aPassportsNameVAK[i])
           IF FILE (mNewName)
              ERASE (mNewName)
           ENDIF
           COPY FILE (mOldName) TO (mNewName)
           mNewName = SUBSTR(aPassportsNameVAK[i],1,6)+'00-'+STRTRAN(STR(j,9),' ','0')+'.txt'
           IF FILE (mNewName)
              ERASE (mNewName)
           COPY FILE (mOldName) TO (mNewName)
           EXTT
        ENDIF
    NEXT
NEXT
CLOSE ALL
LB Warning( 'Процесс переименования файлов завершен успешно!' )
RETURN NIL
```

Правильные наименования файлов паспортов научных специальностей были взяты с сайта ВАК с той же страницы, с которой они скачиваются³³. Это было сделано следующим образом. Паспорта специальностей были выделены блоком, перенесены в MS Excel и записаны в виде DBFIV-файла с именем PassportsNameVAK.dbf, т.е. в виде базы данных, которая используется приведенной выше программой.

Кроме того, с целью формирования обобщенных образов групп специальностей, приведенная выше программа скопировала файлы паспортов специальностей во 2-м стандарте «Эйдос»³⁴. В этом стандарте имена файлов имеют вид: <Наименование класса-#######.txt>, где: ######## – номер реализации объекта, относящегося к классу. В качестве наименований классов были использованы обобщенные до первых двух пар цифр шифры специальностей для и номера примеров, также сгенерированные программой³⁵ (рисунок 3b):

³³ Отметим, что на этой странице они тоже не всегда правильные, т.е. встречаются случаи применения оборванных на полуслове наименований специальностей (например, по специальностям: 05.13.06 Автоматизация и управление технологическими процесс, 05.13.19 Методы и системы защиты информации, информационная).

³⁴ Этот стандарт будет описан чуть ниже.

³⁵ Все это можно было сделать и вручную, но дольше и более трудоемко, чем написать программу и использовать ее для выполнения этой работы.

† Vius	Тип	Размер	Дата	
4 05.12.00-000000138	txt 305.17.00-000000173	bt 🗷 05.23.00-000000216	bt № 06.02.10 Частная зоотехния, технология производства продукто	txt 312.00.00-000000291
4 05.12.00-000000139	bt 405.17.00-000000174	txt 18 05.23.00-000000217	bt ¥ 06.03.00-000000246	bt ¥12.00.00-000000292
¥ 05.12.00-000000140	txt 3605.17.00-000000175	txt 365.25.00-000000218	bt 36.03.00-000000247	tst 🗷 12.00.00-000000293
¥ 05.12.00-000000141	trt 36 05.17.00-000000176	bit 18 05.25,00-000000219	bt ¥ 06.03.00-000000248	bit 12.00.00-000000294
M 05.13.00-000000142	txt 305.18.00-00000177	bd 305.25,00-000000220	bt 1606.04.00-000000249	bt 12.00.00-000000295
¥ 05.13.00-000000143	trt 36 05.18.00-000000178	bd 36.526.00-000000221	bt ¥ 07.00.00-000000250	bit 4 12.00.00-000000296
M 05.13.00-000000144	txt 405.18.00-000000179	txt 305.26.00-00000222	bt 307.00.00-000000251	bit 12.00.00-000000297
4 05.13.00-000000145	bt 405.18.00-00000180	bd 🗷 05.26,00-00000223	bit × 07.00.00-000000252	bt ¥ 12.00.00-000000298
W 05.13.00-000000146	trt 36 05.18.00-000000181	bd × 05.26.00-00000224	bt ¥ 07.00.00-00000253	bit 12,00,00-000000299
4 05.13.00-000000147	bit 405.18.00-000000182	txt (05.27,00-000000225	bit 107.00.00-000000254	bt 12,00,00-00000300
4 05.13.00-000000148	bt # 05.18.00-000000183	txt 305.27.00-000000226	bit # 07.00.00-000000255	bt 12,00,00-00000301
#4 05.13.00-000000149	trt #4 05.18.00-000000184	txt 3 05.27.00-000000227	bit #107.00.00-000000256	bd 4 12.00.00-000000302
M 05.13.00-000000150	bit 14 05.19.00-000000185	txt 14 05.27.00-000000228	bit 1408.00.00-000000257	bt 13,00,00-00000303
#4 05.13.00-000000151	bt # 05.19.00-000000186	bd × 06.01.00-000000229	bd W 08.00.00-000000258	bt × 13,00,00-00000304
34 05.13.00-000000152	tit W 05.19.00-00000187	bd # 06.01.00-00000230	txt 36.00.00-000000259	tst # 13,00,00-00000305
	tit 14 05.19.00-00000188	tid 14(06.01.00-000000231	bd (408,00.00-00000260	txt 34 13,00,00-000000306
	tit % 05.20.00-00000189	trt 14 06.01.00-00000232	bt #608.00.00-000000261	bt #13,00.00-00000307
	tit #105.20.00-00000190	td 4 06.01.00-00000233	bd 14 08.00.00-00000262	txt 34 13.00.00-000000308
	trt 14105.20.00-00000191	trd 180,06.01,00-00000234	bt 1409,00.00-00000263	txt 34 14.01.00-000000309
3 (05.13.1) Управление в социальных и экономических системах 4 (05.13.1) Математическое и программное обеспечение вычислител		tr 14106.01.00-00000235	bt (%)9,00,00-00000264	tit 34 14.01.00-00000310
		trl ¥106.01.01 O6use земледелие	bt 1809.00.00-00000265	txt 34 14.01.00-000000311
	txt 14105.20.02 Электротехнологии и электрооборудование в сельском			
	r txt № 05.20.03 Технологии и средства технического обслуживания в с txt № 05.21.00-000000192	txt ¥ 06.01.02 Мелиорация, рекультивация и охрана земель	bt = 09.00.00-00000266 bt = 09.00.00-00000267	bit 35 14,01.00-000000312
26 05.13.17 Теоретические основы информатики		td: 14 06.01.03 Агрофизика		bit 36 14.01.00-000000313
26 05.13.18 Математическое моделирование, численные методы и ко		txt 36.01.04 Arpownus	txt 109,00,00-00000268	bit 35 14.01.00-000000314
36.13.19 Методы и системы защиты информации, информационная		txt № 06.01.05 Селекция и семеноводство сельскохозяйственных расте	bt 99.00.00-00000269	bit 35 14,01.00-000000315
26 05.13.20 Квантовые методы обработки информации	txt 34 05.22.00-000000195	txt 🗷 06.01.06 Луговодство и лехарственные, эфирно-масличные культ	bd # 09.00.00-000000270	bit 34.01.00-000000316
M05.14.00-000000153	tut 105.22.00-000000196	bd 🗷 06.01.07 Защита растений	txt 309.00.00-000000271	txt 14.01.00-000000317
M05.14.00-00000154	tit ¥ 05.22.00-000000197	trt 1406.02.00-000000236	bt 10.01.00-000000272	bit 4/11/00-000000318
國05.14.00-000000155	tit 105.22.00-000000198	bd = 06.02.00-000000237	bt 10.01.00-000000273	txt 14.01.00-00000319
<u>M</u> 05.14.00-000000156	trt 105.22.00-00000199	txt 14 06.02.00-00000238	bt 10.01.00-000000274	txt 314.01.00-00000320
<u>#</u> 05.14.00-000000157	txt <u>10</u> 05.22.00-00000200	bd <u>=</u> 06.02.00-000000239	bt 4010.01.00-000000275	bt 4.01.00-00000321
M 05.14.00-000000158	tst 305.22.00-000000201	td 106.02.00-00000240	bit 3 10.01.00-000000276	bt 314.01.00-000000322
<u>■</u> 05.14.00-000000159	txt 105.22.00-00000202	bd <u>#</u> 06.02.00-000000241	bit 10.01.00-000000277	bt 14.01.00-00000323
	txt 105.22.00-000000203	bt <u>4</u> 06.02.00-000000242	bt 10.02.00-000000278	tst 35 14.01.00-000000324
■ 05.16.00-000000161	tit 105.23.00-000000204	bit 15 06.02,00-000000243	bt 10.02.00-00000279	bit 14.01.00-00000325
M 05.16.00-000000162	bit 405.23.00-000000205	txt 🗷 06.02.00-000000244	bt 36 10.02.00-000000280	bit 4.01.00-00000326
36 05.16.00-000000163	txt 36 05.23.00-000000206	txt 🗷 06.02.00-000000245	bit № 10.02.00-000000281	bt 14.01.00-000000327
4 05.16.00-00000164	txt 24 05.23.00-000000207	txt 🗷 06.02.01 Диагностика болезней и терапия животных, патология,	bt 10.02.00-00000282	bt 14.01.00-00000328
36 05.16.00-00000165	txt 38 05.23.00-000000208	txt 🗷 06.02.02 Ветеринарная микробиология, вирусология, эпизоотоло	bit 3 10.02.00-000000283	bit 4.01.00-000000329
¥05.16.00-000000166	txt 405.23.00-000000209	txt 🗷 06.02.03 Ветеринарная фармакология с токсикологией	bt 4 10.02.00-00000284	bt 14.01.00-00000330
36 05.16.00-000000167	txt 3605.23.00-000000210	txt 🗷 06.02.04 Ветеринарная хирургия	bt 10.02.00-00000285	txt 14.01.00-000000331
¥ 05.17.00-000000168	txt 1605.23.00-000000211	txt 🗷 06.02.05 Ветеринарная санитария, экология, зоогигиена и вете	bt 10.02.00-00000286	bit 14.01.00-000000332
3 05.17.00-000000169	txt 305.23.00-000000212	bit × 06.02.06 Ветеринарное акушерство и биотехника репродукции жи	bit 3010.02.00-000000287	txt 14.01.00-000000333
¥05.17.00-000000170	tht 305.23.00-000000213	txt 406.02.07 Разведение, селекция и генетика сельскохозяйственны	bt = 12.00.00-000000288	bt 14.01.00-000000334
M 05.17.00-000000171	bit 305.23.00-000000214	bd 4 06.02.08 Кормопроизводство, кормление сельскохозяйственных ж	bit 12.00.00-00000289	tst 14.01.00-000000335
4 05.17.00-000000172	tit 4 05.23.00-000000215	txt ¥ 06.02.09 Звероводство и охотоведение	bt 12.00.00-000000290	bt 414.01.00-000000336

Рисунок 3b. Файлы исходных данных в кодировке DOS OEM 866 (ASCII)

с исправленными именами, а также во 2-м стандарте «Эйдос»

Текстовые файлы исходных данных записываются в поддиректорию: «...\AID_DATA\Inp_data\» в той директории, в которой находится система «Эйдос».

Отметим также, что файлов исходных данных довольно много (429 паспортов специальностей) и они имеют достаточно большой суммарный объем.

Затем с параметрами, показанными на рисунке 4, запускается режим 2.3.2.1 системы «Эйдос», представляющий собой программный интерфейс (API) с внешними данными текстового типа. Этот программный интерфейс не имеет особых ограничений на объем обрабатываемых текстов, т.к. тексты представлены виде файлов, в отличие от интерфейса 2.3.2.2 в котором тексты являются значениями Excel-таблицы и поэтому имеют объем не более 32К. Поэтому в данном случае интерфейс 2.3.2.1 и был использован.

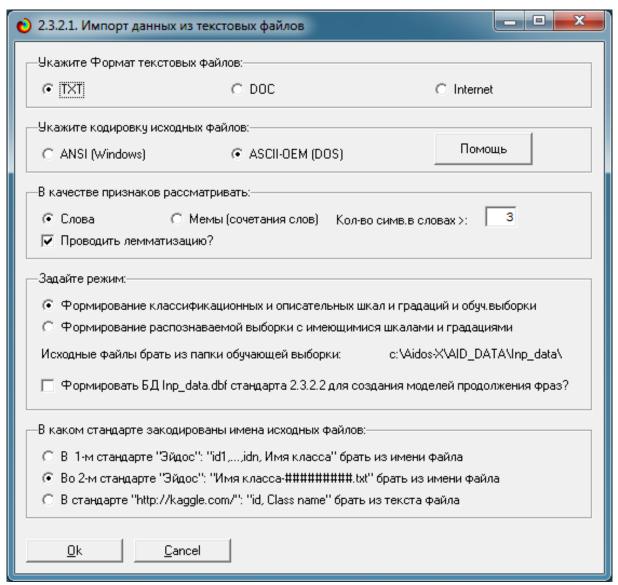


Рисунок 4. Экранные формы программного интерфейса (API) 2.3.2.1 системы «Эйдос» с внешними данными текстового типа

На рисунке 4 приведены реально использованные параметры. Система «Эйдос» обеспечивает рассмотрение в качестве признаков текстов и авторов не только слов, но и их сочетаний по 2, 3 и т.д., некоторые из которых могут оказаться мемами [21] (рисунок 5). Однако в данной работе мы не будем рассматривать этот вариант, т.к. в данном случае в этом нет особой необходимости.

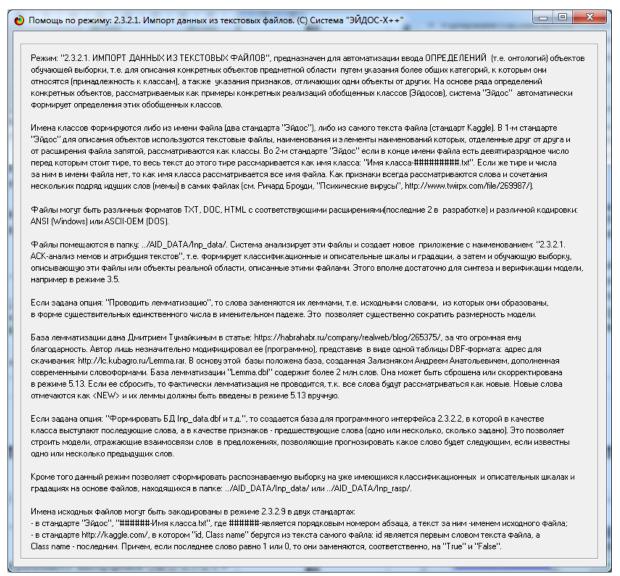


Рисунок 5. Экранные форма HELP программного интерфейса (API) 2.3.2.1 системы «Эйдос» с внешними данными текстового типа

В результате работы режима сформирована одна классификационная шкала (таблица 5) с суммарным количеством классов 82 и одна описательная шкала с числом слов-градаций 11248 (рисунок 6). В качестве градаций описательной шкалы использованы леммы слов исходных текстов (рисунок 3b). Отметим, что без лемматизации количество слов увеличилось бы примерно в два раза.

Система «Эйдос» поддерживает лемматизацию. Леммы получены путем автоматизированного использования словаря академика РАН Андрея Анатольевича Зализняка для лемматизации. Эта база взята из статьи: https://habrahabr.ru/company/realweb/blog/265375/ и преобразована в формат базы данных системы «Эйдос» и входит в полную инсталляцию системы «Эйдос», размещенную на странице сайта автора: http://lc.kubagro.ru/aidos/_Aidos-X.htm. Кроме того эта база размещена на

сайте автора по ссылке: http://lc.kubagro.ru/Lemma.rar (архив имеет размер около 10 Мб, сама база около 217 Мб).

Таблица 5 – Классификационная шкала и ее градации (классы)

u.		іфикационная шкала и се градации (в
	Наименование	
	КЛАСС-01.01.00	
2	КЛАСС-01.02.00	
3	КЛАСС-01.03.00	
4	КЛАСС-01.04.00	
5	КЛАСС-02.00.00	
6	КЛАСС-03.01.00	
	КЛАСС-03.02.00	
8	КЛАСС-03.03.00	
	КЛАСС-05.01.00	
	КЛАСС-05.02.00	
	КЛАСС-05.04.00	
	КЛАСС-05.05.00	
	КЛАСС-05.07.00	
14	КЛАСС-05.08.00	
	КЛАСС-05.09.00 КЛАСС-05.11.00	
	КЛАСС-05.12.00	
18	КЛАСС-05.13.00	
19	KJIACC-05.13.01	Системный анализ, управление и обработка информации
20	KJIACC-05.13.05	Элементы и устройства вычислительной техники и сист
21	КЛАСС-05.13.06	Автоматизация и управление технологическими процесс
22	КЛАСС-05.13.10	Управление в социальных и экономических системах
23	KJIACC-05.13.11	Математическое и программное обеспечение вычислител
24	КЛАСС-05.13.12	Системы автоматизации проектирования (по отраслям)
25	КЛАСС-05.13.15	Вычислительные машины, комплексы и компьютерные сет
26	КЛАСС-05.13.17	Теоретические основы информатики
27	КЛАСС-05.13.18	Математическое моделирование, численные методы и ко
28	КЛАСС-05.13.19	Методы и системы защиты информации, информационная
29	КЛАСС-05.13.20	Квантовые методы обработки информации
30	КЛАСС-05.14.00	
31	КЛАСС-05.16.00	
32	КЛАСС-05.17.00	
	КЛАСС-05.18.00	
34	КЛАСС-05.19.00	
	КЛАСС-05.20.00	
		Технологии и средства механизации сельского хозяйст
37	KDACC-05 20 02	Электротехнологии и электрооборудование в сельском
38	KDACC-05 20 03	Технологии и средства технического обслуживания в с
30	КЛАСС-05.21.00	технологии и средства технического осолуживания в с
40	КЛАСС-05.21.00	
40	КЛАСС-05.22.00 КЛАСС-05.23.00	
	КЛАСС-05.25.00	
	КЛАСС-05.25.00 КЛАСС-05.26.00	
	КЛАСС-05.26.00	
	КЛАСС-05.27.00 КЛАСС-06.01.00	
		Общее земледелие
		Мелиорация, рекультивация и охрана земель
	КЛАСС-06.01.03	
49	КЛАСС-06.01.04	Агрохимия
50	KJIACC-06.01.05	Селекция и семеноводство сельскохозяйственных расте
51	KJIACC-06.01.06	Луговодство и лекарственные, эфирно-масличные культ
52	КЛАСС-06.01.07	Защита растений
53	КЛАСС-06.02.00	
54	KJIACC-06.02.01	Диагностика болезней и терапия животных, патология,
55	КЛАСС-06.02.02	Ветеринарная микробиология, вирусология, эпизоотоло
56	КЛАСС-06.02.03	Ветеринарная фармакология с токсикологией
57	КЛАСС-06.02.04	Ветеринарная хирургия
58	КЛАСС-06.02.05	Ветеринарная санитария, экология, зоогигиена и вете
59	КЛАСС-06.02.06	Ветеринарное акушерство и биотехника репродукции жи
		Разведение, селекция и генетика сельскохозяйственны
61	КЛАСС-06.02.08	Кормопроизводство, кормление сельскохозяйственных ж
62	КЛАСС-06.02.09	Звероводство и охотоведение
63	КЛАСС-06.02.10	Частная зоотехния, технология производства продукто
64	КЛАСС-06.03.00	
65	КЛАСС-06.04.00	
66	КЛАСС-07.00.00	
67	КЛАСС-08.00.00	
	КЛАСС-09.00.00	
69	КЛАСС-10.01.00	
	КЛАСС-10.02.00	
	КЛАСС-12.00.00	
	КЛАСС-13.00.00	
	КЛАСС-14.01.00	
	КЛАСС-14.02.00	
	КЛАСС-14.03.00	
	КЛАСС-14.04.00	
	КЛАСС-17.00.00	
	КЛАСС-17.00.00	
	КЛАСС-19.00.00 КЛАСС-22.00.00	
	КЛАСС-23.00.00	
	КЛАСС-24.00.00	
82	КЛАСС-25.00.00	

При заданных опциях режим 2.3.2.1 в качестве классов используются имена файлов без расширения, а также текст до символа тире: «-» и идущего за ним девятизначного номера файла.

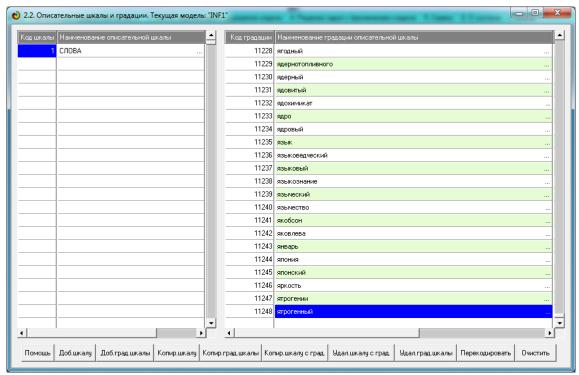


Рисунок 6. Описательные шкалы и градации (фрагмент)

С использованием классификационных и описательных шкал и градаций исходные тексты были закодированы и получена обучающая выборка (рисунок 7):

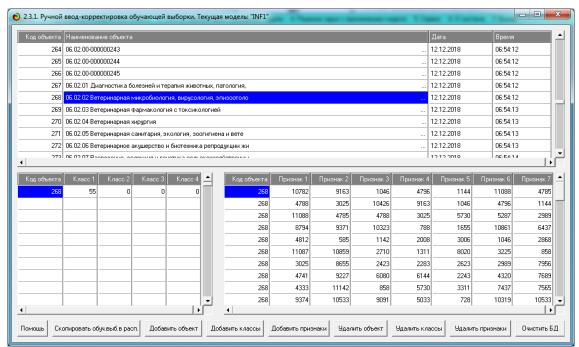


Рисунок 7. Обучающая выборка (фрагмент)

На рисунке 7 приведен лишь фрагмент обучающей выборки, т.к. в нее входит 912 файлов. В верхнем окне на рисунке 7 приведены

наименования объектов обучающей выборки, в левом нижнем окне – коды классов, к которым эти объекты относятся (таблица 5), а в правом нижнем окне – коды признаков, которые есть у этих объектов (рисунок 6).

Обучающая выборка по сути представляет собой нормализованные исходные данные (рисунок 3b).

Таким образом созданы все необходимые и достаточные условия для выполнения следующего этапа ACK-анализа: т.е. для синтеза и верификации моделей.

2.2.2.3.3. Синтез статистических и системно-когнитивных моделей

Синтез и верификация моделей осуществляется в режиме 3.5 системы «Эйдос» (рисунок 8):

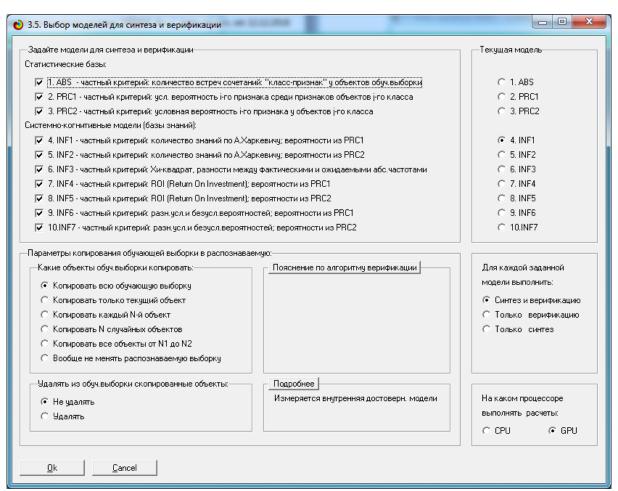


Рисунок 8. Экранная форма режима синтеза и верификации моделей системы «Эйдос»

Обратим внимание на то, что на рисунке 8 в правом нижнем углу окна задана опция: «Расчеты проводить на графическом процессор (GPU)».

Стадия процесса исполнения синтеза и верификации моделей и прогноз времени его окончания отображается на экранной форме (рисунок 9):

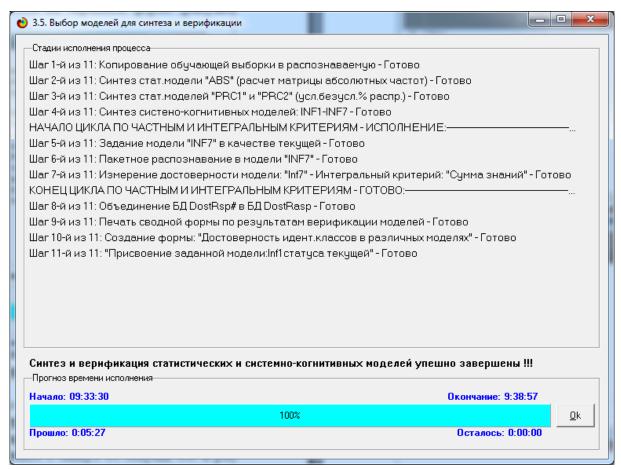


Рисунок 9. Экранная форма с отображением стадия процесса исполнения синтеза и верификации моделей и прогноза времени его окончания

Видно, что весь процесс синтеза и верификации моделей занял 5 минут 27 секунд. Отметим, что при синтезе и верификации моделей использовался графический процессор (GPU) видеокарты. На центральном процессоре (CPU) выполнение этих операций занимает значительно большее время (на некоторых задачах это происходит в десятки, сотни и даже тысячи раз дольше). Таким образом, неграфические вычисления на графических процессорах видеокарты делает возможным обработку реальных текстов за разумное время.

Отметим также, что база лемматизации А.А.Зализняка включает 2097005 слов. В ходе выполнения данной работы она была *дополнена* еще на 11467, представляющих собой научные термины, встретившиеся в паспортах специальностей научных работников ВАК РФ. Эти термины не приводятся в данной работе только из-за ограничения на ее объем.

Фрагменты самих созданных статистических и системно-когнитивных моделей (СК-модели) приведены на рисунках 10-12:

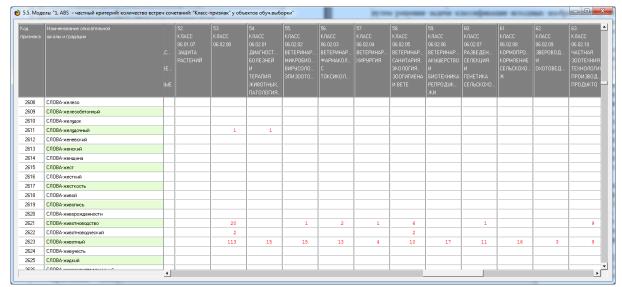


Рисунок 10. Матрица абсолютных частот (фрагмент)

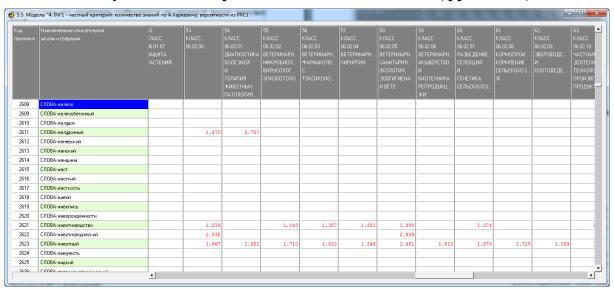


Рисунок 11. Матрица информативностей INF1 (фрагмент)

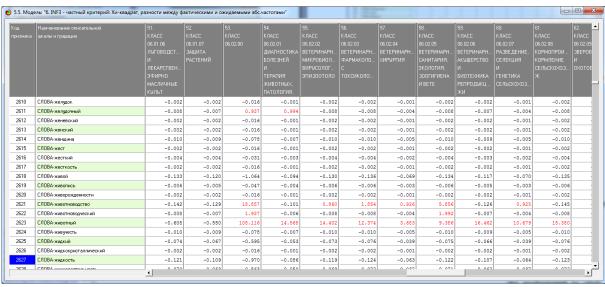


Рисунок 12. Модель INF3 (фрагмент)

2.2.2.3.4. Верификация статистических и системно-когнитивных моделей

Оценка достоверности моделей в системе «Эйдос» осуществляется путем решения задачи классификации исходных изображений по обобщенным образам классов и подсчета количества истинных положительных и отрицательных, а также ложных положительных и отрицательных решений по F-мере Ван Ризбергена и L1- L2-мерам проф.Е.В.Луценко [16].

Классическая количественная мера достоверности моделей: F-мера Ван Ризбергена основана на подсчете суммарного количества верно и ошибочно идентифицированных и не идентифицированных объектов обучающей выборки. В мультиклассовых системах классификации, таких «Эйдос», один объект обучающей система И TOT же распознаваемой выборки может одновременно относится ко многим классам. Соответственно, при синтезе модели его описание используется для формирования обобщенных образов многих классов, к которым он относится. При использовании модели для классификации определяется степень сходства-различия объекта со всеми классами, причем истинноположительным решением может являться принадлежность объекта сразу к нескольким классам. В результате такой классификации получается, что объект не просто правильно или ошибочно относится или не относится к различным классам, как в классической F-мере, но правильно или ошибочно относится или не относится к ним в различной степени.

Однако классическая F-мера не учитывает того, что объект может, одновременно относится ко многим (мультиклассовость) и того, что в результате классификации может быть получена различная степень сходства-различия объекта с классами (нечеткость). Ha многочисленных численных примерах автором установлено, что при истинно-положительных и истинно-отрицательных решениях модуль сходства-различия объекта с классами значительно выше, чем при ложно-положительных и ложно-отрицательных решениях.

Поэтому была предложена L1-мера достоверности моделей [16], учитывающая не просто сам факт истинно или ложно положительного или отрицательного решения, но и степень уверенности классификатора в этих решениях.

Однако при классификации больших данных было обнаружено большое количество ложно-положительных решений с низким уровнем сходства, которые, суммарно могут вносить большой вклад в снижение достоверности модели.

Чтобы преодолеть эту проблему предлагается L2-мера [16], в которой вместо сумм уровней сходства используется средние уровни сходства по различным вариантам классификации.

Таким образом, в системе «Эйдос» применяются достоверности моделей, названые L1-мера и L2-мера, смягчающие и преодолевающие недостатки F-меры. В работе [16] эти меры описаны математически и их применение продемонстрировано на численном примере. В интеллектуальной системе «Эйдос», которая является инструментарием автоматизированного программным системноанализа (АСК-анализ), реализованы когнитивного меры достоверности моделей: F, L1 и L2

В режиме 4.1.3.6 кратко и в режиме 4.1.3.7 более подробно показана достоверность каждой частной модели в соответствии с этими мерами достоверности. В данном случае по L2-мере наивысшую достоверность имеет модель INF3 (рисунок 13):

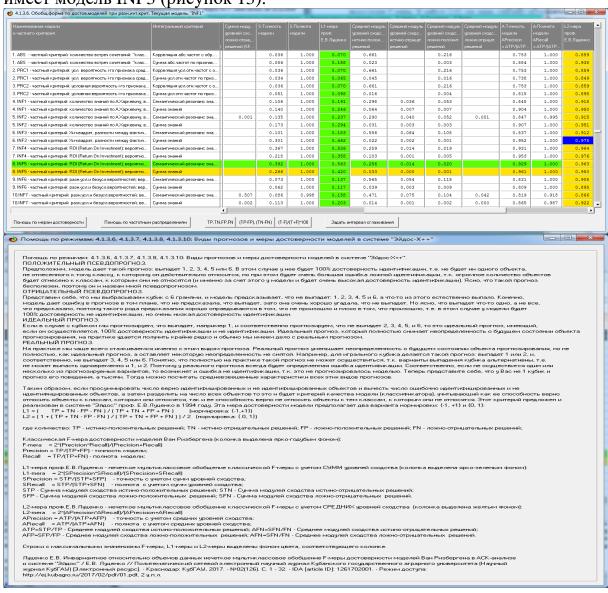


Рисунок 13. Экранная форма с информацией о достоверности моделей по F-критерию Ван Ризбергена и L1- и L2-криетриям проф.Е.В.Луценко [16]

Из рисунка 13 мы видим, что в данном интеллектуальном приложении по критерию L2 наиболее достоверной (0,980) является модель INF5 с интегральным критерием «сумма знаний» (ROI). Не очень уступает наиболее достоверной модели и СК-модель INF3 (частный критерий, аналогичный хи-квадрат) (0,975), которую мы и примем для рассмотрения.

На рисунке 14 приведены частные распределения уровней сходства и различия истинных и ложных положительных и отрицательных решений в модели INF3 (получены в режиме 3.4):

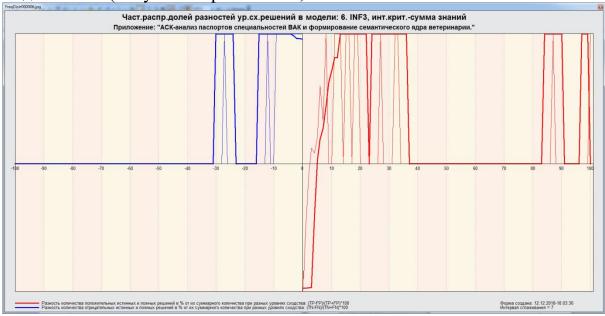


Рисунок 14. Частотные распределения уровней сходства и различия истинных и ложных положительных и отрицательных решений в модели INF3

Из рисунка 14 мы видим, что отрицательные решения (т.е. решения о непринадлежности объекта к классу) в данной модели всегда истинные, а положительные решения (т.е. о принадлежности объекта у классу) есть и истинные, и ложные, причем все ложные решения наблюдаются при невысоких уровнях сходства (примерно до 5%), а более высоких уровнях сходства наблюдаются только истинно-положительные решения. Это позволяет надежно и с высокой достоверностью решать задачи системной классификации и идентификации текстов.

Это очень хорошие и разумные результаты.

2.2.2.3.5. Повышение качества модели

Обратимся к режиму 3.7.5. Данный режим показывает Паретозависимость суммарной дифференцирующей мощности модели от числа градаций описательных шкал, рассортированных в порядке убывания их селективной силы или значимости (рисунок 15).

В качестве значимости значения фактора для решения различных задач в системно-когнитивной модели (СК-модель) в АСК-анализе и

системе «Эйдос» принята *вариабельность* значений фактора по классам в матрице модели (таблица 4).

В качестве значимости фактора для решения различных задач в системно-когнитивной модели (СК-модель) в АСК-анализе и системе «Эйдос» принята средняя *значимость значений данного фактора* по классам в матрице модели (таблица 4).

Существует много количественных мер вариабельности (изменчивости), но в данном случае принято использовать среднеквадратичное отклонение 36 .

Вариабельность значения фактора по классам выбрана в качестве значимости этого значения фактора потому, что чем выше эта вариабельность, тем лучше позволяет это значение фактора разделить (различить) классы. Если вариабельность значения фактора равна нулю, то данное значение фактора вообще является бесполезным для решения задачи идентификации и других задач.

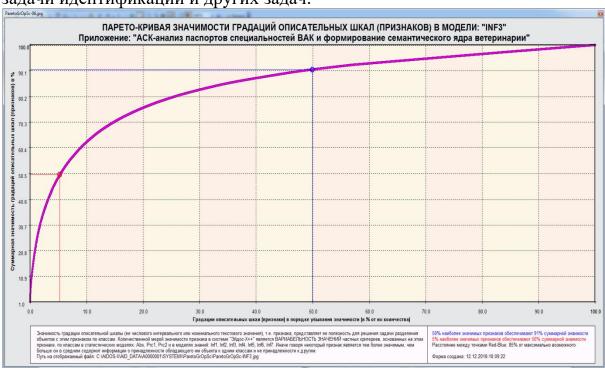


Рисунок 15. Парето-зависимость суммарной дифференцирующей мощности модели от числа градаций описательных шкал, рассортированных в порядке убывания их селективной силы в СКмодели INF3

На основе рисунка 15 и соответствующих таблиц, которые здесь не приводятся из-за ограниченности объема статьи, можно обоснованно

 $^{^{36}}$ Не путать со стандартным отклонением, которое является одним из параметров распределения Гаусса.

сделать выводы о том, что, например, в наиболее достоверной модели INF3:

- 1. 50% наиболее значимых градаций описательных шкал обеспечивают 91% суммарной селективной мощности модели (точка BLUE).
- 2. 50% суммарной селективной мощности модели обеспечивается 5% наиболее значимых градаций описательных шкал (точка RED).
- 3. Число градаций описательных шкал может быть существенно сокращено без особой потери качества модели путем удаления из модели малозначимых градаций. При этом размерность модели существенно сократится и ее быстродействие соответственно возрастет.

2.2.2.3.6. Выбор наиболее достоверной модели и присвоение ей статуса текущей

В соответствии со схемой обработки данных, информации и знаний в системе «Эйдос» (рисунок 2), присвоим наиболее достоверной СК-модели INF3 статус текущей модели. Для это запустим режим 5.6 с параметрами, приведенными на экранной форме (рисунок 16):

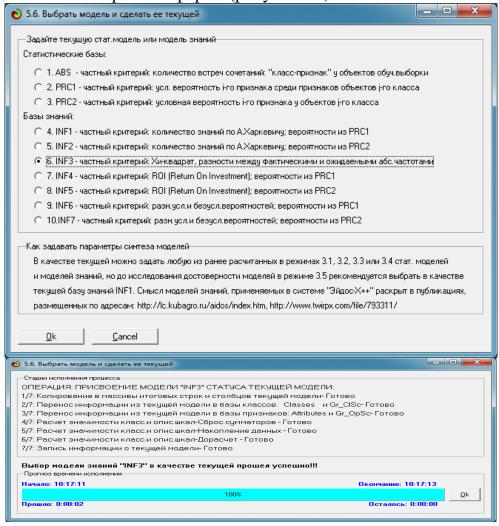


Рисунок 16. Экранные формы придания модели статуса текущей

3.2.3. Результаты и обсуждение

3.2.3.1. Диагностика (классификация, распознавание, идентификация) (задача 3)

3.2.3.1.1. Верификация моделей путем распознавания обучающей выборки

Решим задачу идентификации обучающей выборки в наиболее достоверной модели на GPU. Для этого запустим режим 4.1.2 (рисунок 17):

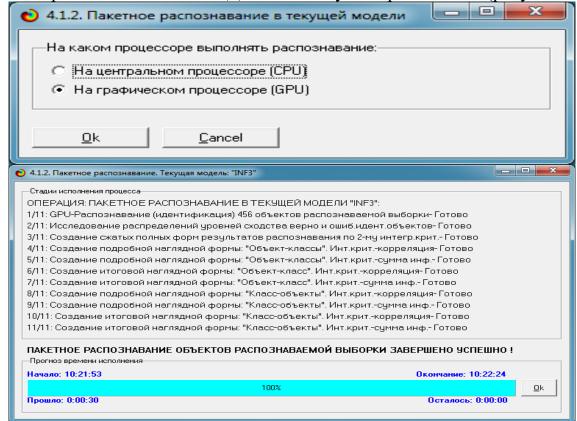


Рисунок 17. Экранная форма отображения процесса решения задачи идентификации в наиболее достоверной модели

Видно, что идентификация 82 текстов в ранее созданных моделях заняла 30 секунд. Отметим, что 99,999% этого времени заняла не сама идентификация на GPU, а создание 10 выходных форм на основе результатов этой идентификации. Эти формы отражают результаты идентификации в различных разрезах и обобщениях:

4.1.3.1. Подробно наглядно: "Объект - классы"

4.1.3.2. Подробно наглядно: "Класс - объекты"

4.1.3.3. Итоги наглядно: "Объект - класс"

4.1.3.4. Итоги наглядно: "Класс - объект"

4.1.3.5. Подробно сжато: "Объекты - классы"

4.1.3.6. Обобщ форма по достов.моделей при разных интегральных крит.

4.1.3.7. Обобщ стат.анализ результатов идент. по моделям и инт.крит.

4.1.3.8. Стат.анализ результ. идент. по классам, моделям и инт.крит.

4.1.3.9. Достоверность идент.объектов при разных моделях и инт.крит.

4.1.3.10. Достоверность идент.классов при разных моделях и инт.крит.

4.1.3.11. Распределения уровн.сходства при разных моделях и инт.крит.

4.1.3.12. Объединение в одной БД строк по самым достоверным моделям

Приведем две формы из этих 10: 4.1.3.1 и 4.1.3.2 (рисунок 18):

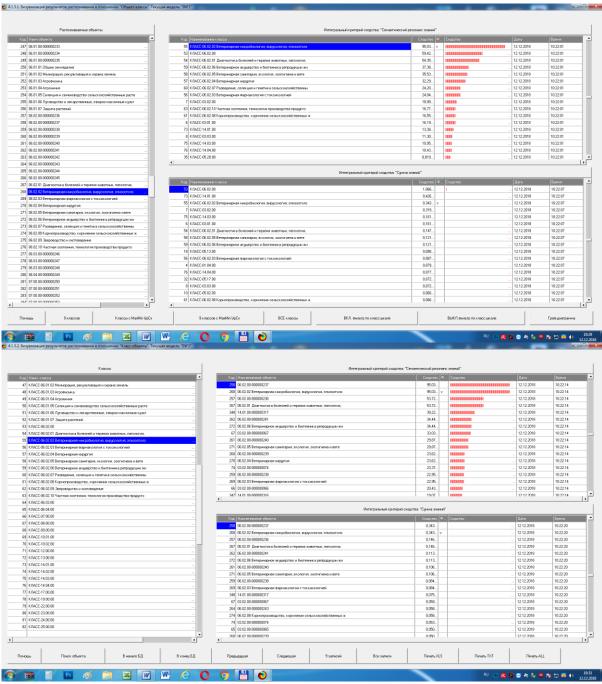


Рисунок 18. Выходные формы по результатам верификации моделей

Из рисунка 18 видно, что результаты атрибуции текстов очень хорошие и полностью соответствуют экспертным ожиданиям, основанным на интуиции, опыте и профессиональной компетенции.

3.2.3.1.2. Классификация текстов, не использовавшихся при синтезе и верификации моделей

Возникает вопрос о том, как же сравнить с классами тексты, которые не использовались в обучающей выборке для синтеза статистических и СК-моделей. Для этого необходимо записать эти тексты в папку: c:\Aidos-X\AID_DATA\Inp_rasp, а затем в программном интерфейсе 2.3.2.1 ввести эти тексты в распознаваемую выборку, задав соответствующую опцию (рисунок 19).

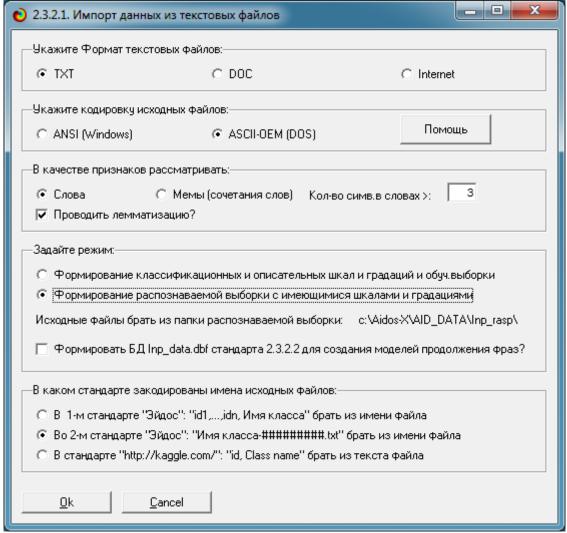


Рисунок 19. Экранная форма программного интерфейса с опциями для ввода текстов в распознаваемую выборку

Мы для нашего примера скопировали все текстовые файлы из папки обучающей выборки в папку распознаваемой выборки, а также скопировали в нее же три текстовых файла в кодировке DOS OEM 866

(переносы выключены) – учебники по механизации [23], агрономии [24] и ветеринарии [25].

В результате работы программного интерфейса 2.3.2.1 получена распознаваемая выборка, фрагмент которой приведен на рисунке 20. Сам процесс распознавания распознаваемой выборки, состоящей из обучающей выборки и трех учебников [23, 24, 25], занял 31 секунду (рисунок 21), что всего на 1 секунду больше, чем было без учебников (рисунок 17).

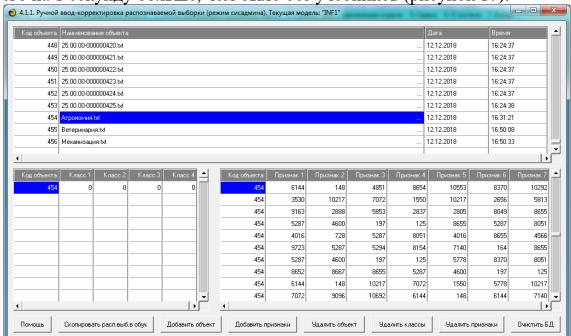


Рисунок 20. Экранная форма программного интерфейса с опциями для ввода текстов в распознаваемую выборку (фрагмент)

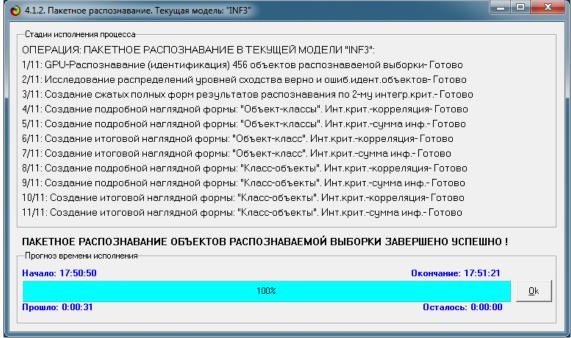


Рисунок 21. Экранная форма отображения процесса решения задачи идентификации в наиболее достоверной модели

Сами результаты идентификации учебных пособий по классам, сформированным на основе паспортов специальностей и групп специальностей научных работников ВАК РФ, приведен на рисунках 22, 23 и 24:

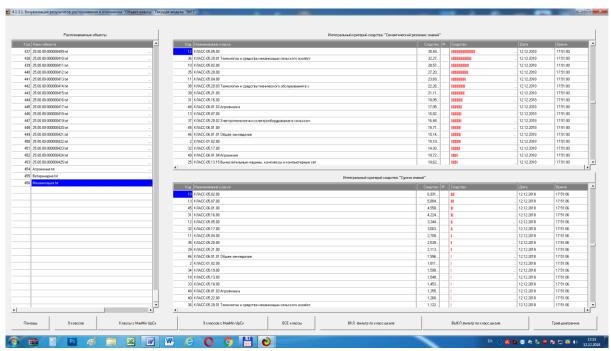


Рисунок 22. Результат идентификации учебного пособия по механизации

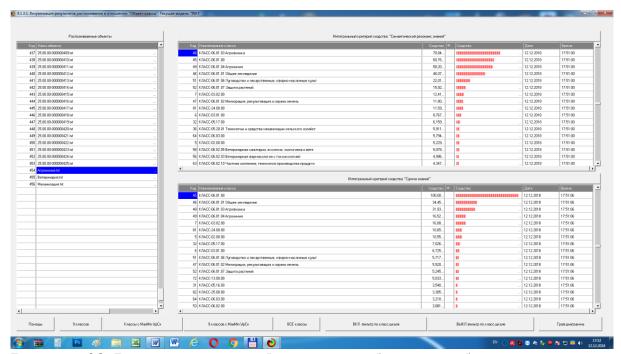


Рисунок 23. Результат идентификации учебного пособия по агрономии

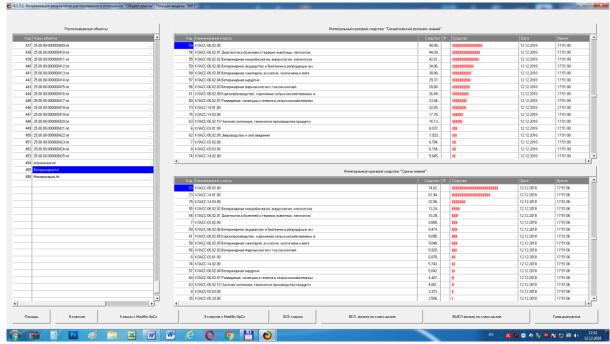


Рисунок 24. Результат идентификации учебного пособия по ветеринарии

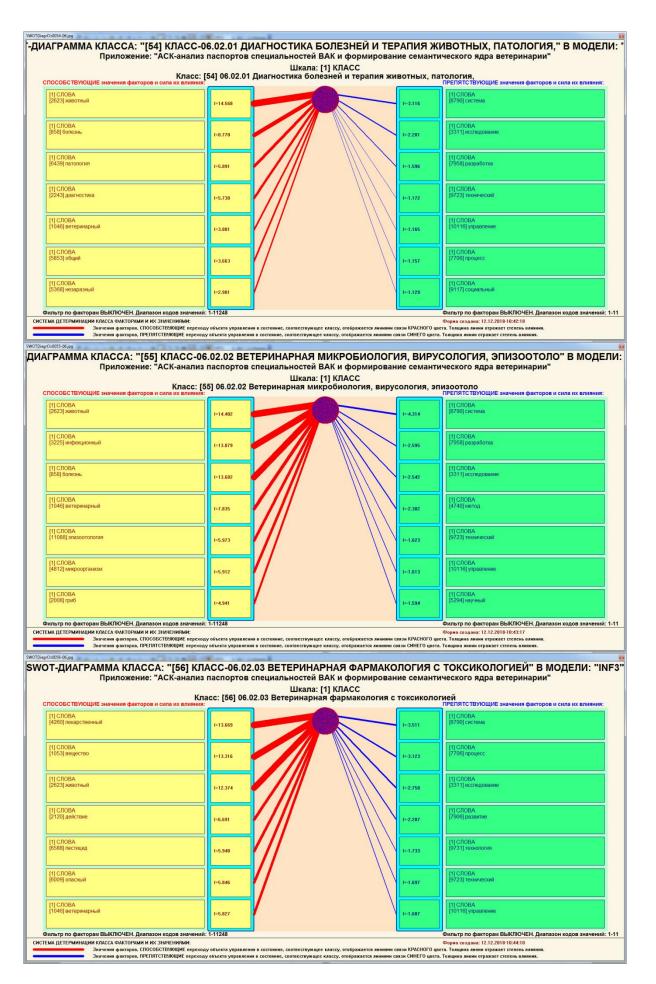
Из анализа форм по результатам классификации учебных пособий [23, 24, 25], представленных на рисунках 22, 23 и 24, мы можем обоснованно сделать следующие выводы:

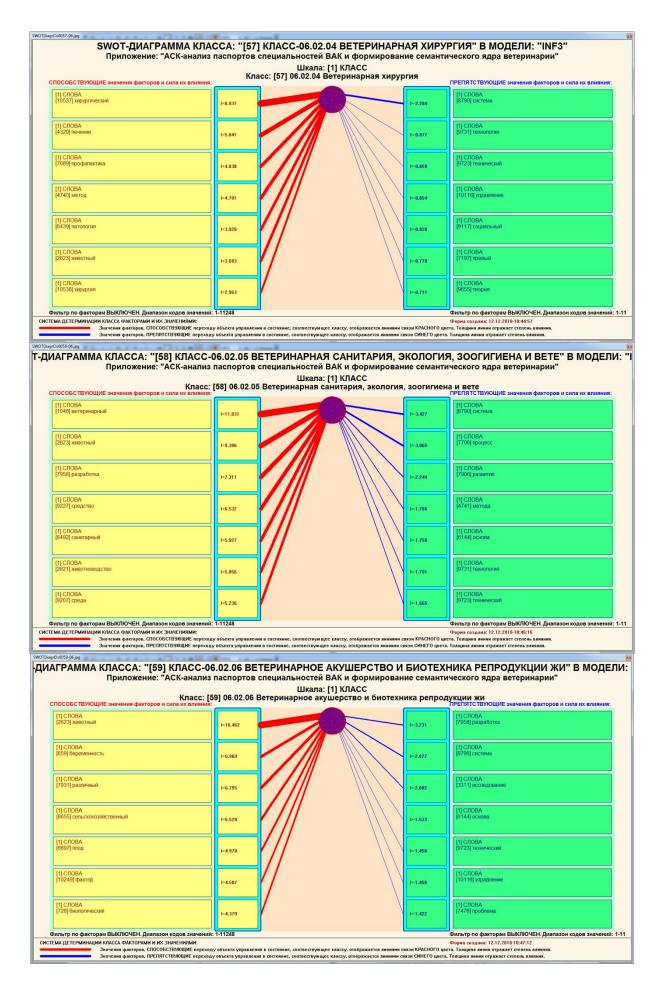
- 1. Метод АСК-анализа и его программный инструментарий, в качестве которого в настоящее время выступает интеллектуальная система «Эйдос», обеспечивают качественное решение задачи автоматической классификации текстов по направлениям науки.
- 2. Видно, что каждое учебное пособие в несколько различной степени относится к нескольким смежным направлениям науки. Так и должно быть, т.к. по своему содержанию каждое учебное пособие охватывает группу специальностей, шифры которых совпадают по первым четырем цифрам.

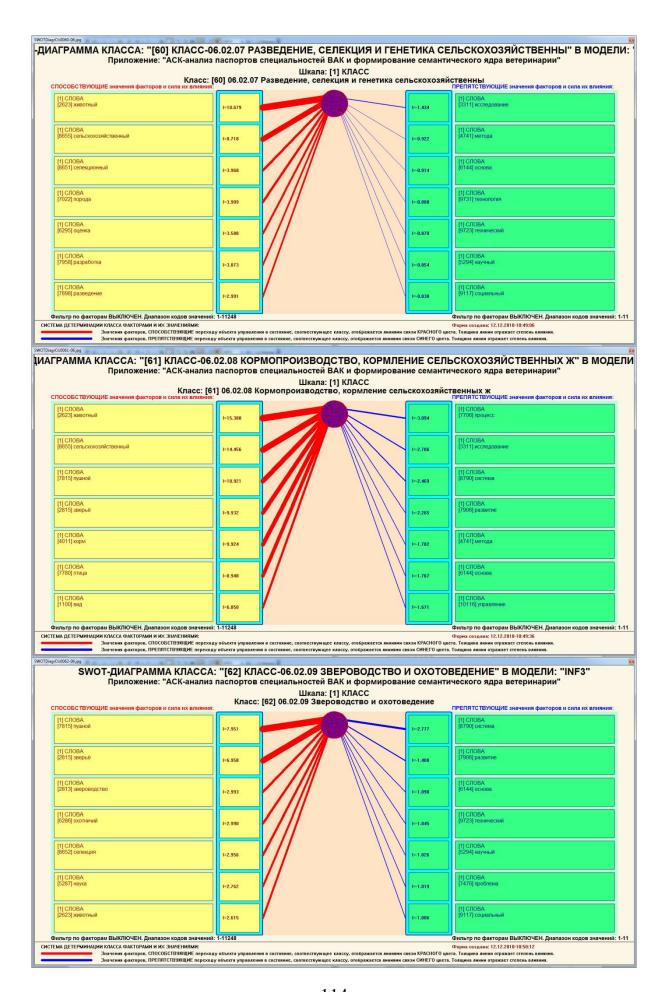
Аналогично могут быть созданы семантические ядра и по другим способам классификации текстов и по ним также могут решаться задачи классификации.

3.2.3.2. Поддержка принятия решений (задача 3)

При принятии решений определяется сила и направления влияния факторов на принадлежность состояний объекта моделирования к тем или иным классам. По сути это решение задачи SWOT-анализа [17]. В системе «Эйдос» в режиме 4.4.8 поддерживается решение этой задачи. При этом выявляется семантическое ядро и семантическое антиядро (термин авт.) заданного класса. На рисунках 25 приведены семантические ядра и антиядра всех ветеринарных специальностей, а также ветеринарии в целом.







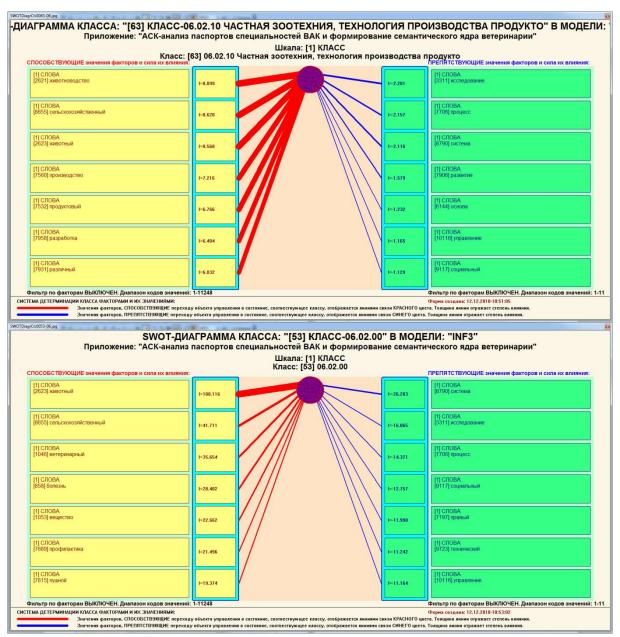


Рисунок 25. Экранные формы решения задачи SWOT-анализа, содержащие семантические ядра и антиядра по всем ветеринарным специальностям ВАК РФ и ветеринарии в целом (2 последние формы)

Выходные формы, приведенные на рисунках 25, как говорят «интуитивно понятны», т.е. не требуют особых комментариев. Отметим лишь, что информация о семантических ядрах и антиядрах ветеринарных специальностей и ветеринарии в целом может быть приведена не только в приведенных, но и во многих других табличных и графических формах, которые в данной работе не приводятся только из-за ограниченности ее размера. В частности в этих формах может быть выведена значительно более полная информация (в т.ч. вообще вся имеющая в модели). Подобная подробная информация содержится базе данных, расположенной c:\Aidosследующему пути: ПО

$X AID_DATA A0000001 System SWOTCls 0053 Inf 3. DBF^{37}.$

семантическом ядре приведены слова, ранжированные в порядке убывания степени их характерности для обобщенного лингвистического образа класса «06.02.00-Ветеринария», а в семантическом антиядре — слова в приведенные в порядке убывания степени не характерности.

В

Фрагмент этой базы приведен в таблице 6:

Таблица 6 — Семантическое ядро и семантическое антиядро обобщенного лингвистического образа класса «06.02.00-Ветеринария» (фрагмент)

(фрагмент)								
		Семантическо	е ядро			Семантиче	ское антиядро	
Nº	Код1	Наименование-1	Значимость	Значимость в %	Код2	Наименование-2	Значимость	Значимость в %
1	2623	животный	108,1163788	6,7682193	8790	система	-26,2027434	1,6403254
2	8655	сельскохозяйственный	41,7111816	2,6111717	3311	исследование	-16,0654067	1,0057151
3	1046	ветеринарный	35,6538734	2,2319767	7706	процесс	-14,3711308	0,8996513
4	858	болезнь	28,4016647	1,7779794	9117	социальный	-12,7568864	0,7985975
5	1053	вещество	22,6624565	1,4186979	7197	правый	-11,9899072	0,7505836
6	7689	профилактика	21,4955807	1,3456500	9723	технический	-11,2421168	0,7037710
7 8	7815	пушной	19,3738956	1,2128299	10116	управление	-11,1638545	0,6988717
9	3225 2621	инфекционный	19,0138855 18,8573589	1,1902928 1,1804940	2233 9655	деятельность	-10,9411820 -9,9568347	0,6849321 0,6233107
10	4011	животноводство корм	17,4052010	1,0895872	7906	теория	-9,8439874	0,6162463
11	728	биологический	17,3650570	1,0870741	10864	развитие экономический	-8,3463820	0,5224943
12	2815	зверьё	16,4678116	1,0309054	6922	политический	-8,2176255	0,5144339
13	7931	различный	15,0588179	0,9427007	7476	проблема	-7,9134120	0,4953898
14	7958	разработка	14,6669884	0,9181717	3239	информационный	-7,6697840	0,4801383
15	5287	наука	14,6374512	0,9163226	9729	технологический	-7,5758685	0,4742591
16	4320	лечение	14,5425386	0,9103809	7170	правовой	-7,2958888	0,4567320
17	6439	патология	13,7634430	0,8616086	6144	основа	-6,9151801	0,4328992
18	4260	лекарственный	13,4173174	0,8399407	3334	история	-6,8715004	0,4301648
19	4740	метод	13,2168550	0,8273915	4551	математический	-6,8401951	0,4282050
20	7780	птица	12,5930319	0,7883394	306	анализ	-6,6402468	0,4156880
21	7532	продуктовый	11,3547068	0,7108187	4552	материал	-6,6228265	0,4145975
22	4812	микроорганизм	11,2799797	0,7061407	9003	создание	-6,3446144	0,3971810
23	2243	диагностика	10,9477386	0,6853420	4181	культура	-6,2923535	0,3939094
24	6009	опасный	8,7947483	0,5505621	6147	основный	-6,1671328	0,3860704
25 26	4013 6437	кормление	8,7182531	0,5457734	10326	физический	-5,7619323	0,3607044
27	7531	патогенный	8,5617266 8,5617266	0,5359747 0,5359747	8592 6081	СВОЙСТВО	-5,4488796 -5,4211101	0,3411069 0,3393685
28	1100	продуктивный вид	8,5180473	0,5332403	10319	организация физик	-5,3236588	0,3332679
29	8959	совершенствование	8,4867420	0,5312806	5712	обеспечение	-5,2802361	0,3305496
30	8492	санитарный	8,4208527	0,5271558	8320	решение	-5,0732168	0,3175899
31	9227	средство	8,3231449	0,5210392	11235	язык	-5,0401439	0,3155195
32	3527	качества	8,2416029	0,5159346	6232	отношение	-5,0401439	0,3155195
33	5132	надзор	8,2173691	0,5144175	8975	современный	-4,9028057	0,3069219
34	3298	использование	8,0936584	0,5066730	7347	прибор	-4,7427439	0,2969019
35	10537	хирургический	8,0295372	0,5026590	10893	эксплуатация	-4,6175231	0,2890629
36	2990	изучение	7,8709865	0,4927335	7540	проектирование	-4,4471120	0,2783949
37	5853	общий	7,1964130	0,4505044	10415	формирование	-4,4019216	0,2755660
38	5368	незаразный	6,7808633	0,4244904	7753	психологический	-4,3983864	0,2753446
39	659	беременность	6,7182527	0,4205709	9487	сфера	-4,2575129	0,2665258
40 41	2423	домашний	6,6556425	0,4166515	3333 9722	исторический	-4,2575129	0,2665258
42	7565 6588	происхождение	6,6382222 6,5930319	0,4155609 0,4127320	3240	техника	-4,2123226 -4,1949023	0,2636968 0,2626063
43	2120	пестицид	6,5894966	0,4125107	5773	информация	-4,0714492	0,2548780
43	8652	действие селекция	6,4365058	0,4123107	10500	обработка функционирование	-4,0696815	0,2547673
45	2653	заболевание	6,3842444	0,3996616	5294	научный	-4,0073279	0,2508639
46	1656	генетический	6,1686435	0,3861647	10619	цель	-3,9618808	0,2480188
47	10322	физиологический	6,1547585	0,3852955	1060	взаимодействие	-3,9462282	0,2470390
48	3067	инвазионный	5,8121686	0,3638490	11237	языковый	-3,8975026	0,2439887
49	4791	микотоксинов	5,8121686	0,3638490	4566	машина	-3,8836175	0,2431195
50	11088	эпизоотология	5,7808633	0,3618892	5778	образование	-3,8035869	0,2381094
51	9795	токсичный	5,7495580	0,3599295	9371	структура	-3,7758171	0,2363710
52	5730	область	5,7351594	0,3590281	10189	устройство	-3,6940185	0,2312503
53	6295	оценка	5,7232990	0,3582856	4600	международный	-3,6644808	0,2294012
54	2989	изучающий	5,6695271	0,3549194	9163	специальность	-3,6425284	0,2280270
55	7022	порода	5,6069169	0,3509999	7004	понятие	-3,6001028	0,2253711
56 57	7956	разрабатывающий	5,5930319	0,3501307	1983 404	гражданский	-3,5844501 3,5844501	0,2243912 0,2243912
57 58	10688 7530	частный	5,3721275 5,2017164	0,3363018 0,3256339	2954	аппарат	-3,5844501 -3,5374922	0,2214516
59	2670	продуктивность загрязнение	5,0451903	0,3158352	8141	изделие регулирование	-3,4296917	0,2147031
60	6627	питание	4,9669271	0,3109358	4746	методология	-3,4122712	0,2136126
61	9191	способность	4,9356217	0,3089760	10512	характеристика	-3,3809661	0,2116528
62	4012	корма	4,8434739	0,3032075	2679	задача	-3,3809661	0,2116528
63	4142	кроликовый	4,8434739	0,3032075	8602	СВЯЗЬ	-3,3653134	0,2106729
64	5763	обоснование	4,8207502	0,3017849	3954	контроль	-3,3496608	0,2096931
65	8834	СКОТ	4,8121686	0,3012477	9538	также	-3,3236588	0,2080653
66	6697	плод	4,7965159	0,3002678	2064	дать	-3,2749332	0,2050150

 37 Отметим, что dbf-файлы нормально открываются в MS Excel, в который встроен соответствующий конвертер.

.

Полностью данная таблица не приводится, т.к. в ней 10624 строки. Поэтому ниже приводятся лишь слова из семантического ядра для группы специальностей: «06.02.00-Ветеринария», в порядке убывания степени их характерности, но без ее указания.

Слова ядра «06.02.00-Ветеринария» в порядке убывания степени

их характерности: животный, сельскохозяйственный, ветеринарный, болезнь, вещество, профилактика, пушной, инфекционный, животноводство, корм, биологический, различный, разработка, наука, лечение, патология, лекарственный, метод, птица, продуктовый, микроорганизм, диагностика, опасный, кормление, патогенный, продуктивный, совершенствование, санитарный, средство, качества, надзор, использование, хирургический, изучение, общий, незаразный, беременность, домашний, происхождение, пестицид, действие, заболевание, генетический, физиологический, инвазионный, микотоксинов, эпизоотология, токсичный, область, оценка, изучающий, порода, разрабатывающий, частный, продуктивность, загрязнение, питание, способность, корма, кроликовый, обоснование, скот, плод, мероприятие, хозяйственный, генетик, гриб, обмен, селекционный, вирус, морфология, патогенез, значение, норма, терапия, продукция, производство, организм, содержание, метода, способ, нутрия, воспроизводительный, разведение, племенной, рацион, остаток, объектовый, загрязнитель, химический, питательный, токсичность, активный, воспроизводственный, кормовой, иммунология, окружающий, фармакология, инфекция, популяция, обосновать, воспроизводство, эффективный, мониторинг. хирургия, борьба, морфологический, новый, проявление, направленный, биотехника, грызун, звероводство, зверь, зоокультуру, кролик, нейроэндокринный, ондатра, осеменение, рецепт, самка, сперма, фитотоксин, эпизоотологический, вводимый, обеззараживание, индикация, комбикормовый, охотничий, отравление, репродукция, шифр, бактерия, производитель, гинекологический, токсикология, разрабатываем, формула, биотехнологический, которая, семиотика, разный, метаболизм, заключаться, иммунологический, помещение, затрата, вода, физиология, фактор, регуляция, клинический, получение, существовать, продукт, антропогенный, возникновение, предупреждение, экспертиза, заниматься, другой, клетка, народный, проведение, бифенил, гамета, домик, енотовидный, животноводческий, зоогигиена, зоогигиенический, зоотехния, иммуноморфологический, контаминант, контаминированных, кормопроизводство, лисица, мелкий, микотоксикологией, нозология, норка, охота, полибромированных, охотоведение, песец, поголовье, пчела, раздача, санитария, скармливание, собака, соболий, стимуляция, сурок, тератогенный, хлордиоксинов, хорёк, шиншилла, шкурка, эмбрион, эмбриотоксического, искусственный, витаминный, генотип, иммунодефицит, клещи, культивирование, мать, выздоровление, единицу, микология, мутагенный, этология, элемент, акушерство, резистентность, этиология, вызываемый, эндогенный, микробиология, онкологический, онкология, вирусология, меховой, наименьший, репродуктивный, роды, тканевый, трансплантация, экзогенный, диагностик, повышение, генетика, типовой, травматизм, ветеринария, доза, побочный, показание, уровень, исход, ткань, фармакологический, обеспечивающий, народнохозяйственный, объективный, теоретический, влиять, введение, естественного, радионуклид, численность, насекомые, увеличение, изучаем, крупный, ликвидация, потребность, вопрос, функция, ведение, взаимосвязь, технология, нарушение, специфический, линия, выращивание, клеточный, ранний, практика, связать, состоять, регион, патологический, течение, ценность, биохимический, идентификация, биология, дифференциальный, отходы, минеральный, этой, снижение, распространение, причина, акарицидное, аллергенный, антибактериальный, антигельминтный, антидотный, антипротозойное, биотехния, биоцид, бэтаагонистов, вакцинология, внедряем, водопоения, воспаление, всасывание, выгул, выживаемость, годный, гонадотоксического, дичеразведение, зооантропоноз, зоотическая, зоотических, инсектицидный, канцерогенного, кардио, кастрация, кишечные, коза, консервированный, кормоприготовления, корова, кратность, кросс, лошадь, маститый, мигрировать, микотоксикозов, микотоксикологический, микотоксинами,

микробиального, наследуемость, млекопитающие, молодняк, молоко, накапливаться, нефропатология, нитрит, овца, окраска, онкогенез, опушение, отапливаемый, охотохозяйственный, очаговость, пат, питательность, плотоядный, повторяемость, поение, полихлорированных, полноценность, полноценный, полуавтоматического, премикс, препаративного, противоинфекционный, противопоказание, противоэпизоотический, раненый, респирационные, рогатый, родентицидное, родовспоможение, свинья, свободноживущий, серология, серопрофилактика, серотерапия, синдроматика, скрещивание, стадо, стильбенов, токсигенных, стрессовый, сырой, токсигенными, токсикоз, тормозящий, убойный, фармакодинамика, фетоплацентарная, чистопородный, шелкопряд, эпизоотический, ядовитый, антивирусный, благополучие, вакцина, вакцинация, вскармливание, вызывающий. гепатогенный, дезактивация, дикий, иммунопатология, импортный, наибольший, нейроактивный, оправдать, патоморфологических, подбор, полиморфизм, породоиспытание, сопряженность, стероидный, стимулирующий, успех, штамм, экстерьер, этиологию, анатом, гистология, гистохимических, диетический, желудочный, зоотехнический, витамин, иммунопатологический, исчезать, макроорганизм, микроскопический, неразрывный, патогенность, предрасположенность, производимый, пушнина, семейство, трансмиссивный, увеличить, среда, акклиматизация, биопрепарат, высококачественный, генофонд, детоксикация, концентрат, обследование, патофизиология, переливание, переносчик, приспособительный, условный, эндокринология, качеством, балансовый, белковый, гематологический, качественной, конъюнктура, недостаточность, постройка, респираторный, род, эмбриология, устойчивость, аутоиммунный, допустимый, исследующий, картирования, максимальный, моделировании, новообразование, обезвреживание, оплата, организменный, получаемый, пригодность, естественной, кадастра, морфогенез, накопление, нитрат, регенерация, рецептура, тога, физиотерапия, опыт, период, выведение, загрязненный, систематика, вновь, интродукция, механизированный, посредство, себестоимостью, условие, белок, перерабатывающий, популяционный, радиация, биосинтез, вариант, выяснение, диспансеризация, расстройство, результативность, хозяйствование, объект, добавка, совместимость, показатель, аномалия, гормон, кожевенный, этого, возможность, обусловленность, потеря, промысловый, выброс, размножение, уход, фонд, эндокринный, возбудитель, прогнозирование, мера, повышенный, рудный, биохимия, надежный, норматив, обусловить, видовой, восстановительный, защитный, изменчивость, иммунитет, отбор, поражение, профилактический, сохранение, военный, деть, природопользование, прогрессивный, размер, травма, характеризующий, актуальный, кровь, нормирование, половый, устранение, обобщение, представляем, протекать, радиационный, установление, структуральный, наследственный, низкий, автоматический, наиболее, лучевой, ними, применительно, чистый, подготовка, изыскание, позволяющий, расширение, инструментальный, первичный, полевой, препарат, зона, способствовать, динамик, собой. этих. выделение, операция, характер, единиц, специализированный, микроавтобус, открыть, испытание, механизация, тканьё, принцип, больные, используемый, страна, возраст, иметь, коррекция, определяющий, организационный, служба, адаптация, лабораторный, полый, рост, вредный, зависимость, конструирование, реабилитация, устойчивый, рациональный, специальный, отечественный, соответствующий, поверхность, структурный, применение, дисциплина, планирование, контрольный, растительный, стандартизация, тело.

Слова антиядра «06.02.00-Ветеринария» не приводятся, т.к. их 10623. Отметим, что система «Эйдос» обеспечивала решение этой задачи (формирование семантического ядра и антиядра) всегда, даже в самых ранних версиях. Первый акт внедрения системы «Эйдос», где об этом упоминается в явном виде, датируется 1987 годом.

Y T B E P H A A D Заведующий Краснодарским сектором ИСИ АН СССР, к. ф. н. А. А. Хагуров 1987г.

утверждаю Директор Северо-Кавказского филиала ВНИЦ "АИУС-агроресурсы", к.э.н. Э.М. Трахов 1987r.

Настоящий акт составлен комиссией в составе: Кириченко М.М., Ляшко Г.А., Самсонов Г.А., Коренец В.И., Луценко Е.В. в том, что в соответствии с договором о научно-техническом сотрудничестве между Сезеро-Кавказским филиалом ВНИЦ "АИУС-агроресурсы" и Краснодарским сектором Института социологических исследований АН СССР Североботы:

- осуществлена постановка задачи: "Обработка на ЭВМ социологических анкет Крайагропрома;
- разработаны математическая модель и программное обеспечение подсистемы распознавания образов, позволяющие решать данную задачу в среде персональной технологической системы ВЕГА-М;
- на профессиональной персональной ЭВМ "Искра-226" осуществлены расчёты по задаче в объёме:

Входная информация составила 425 анкет по 9-ти предприятиям. Выходная информация — 4 вида выходных форм объёмом 90 листов формата А3 и 20 листов формата А4 содержит:

- процентное распределение ответов в разрезе по социальным типам корреспондентов;

- распределение информативностей признаков (в битах) для распозна-

вания социальных типов корреспондентов;

- позитивные и негативные информационные портреты 30-ти социальных типов на языке 212 признаков;

- обобщённая характеристика информативности признаков для выбора такого минимального набора признаков, который содержит максимум информации о распознаваемых объектах (оптимизация анкет).

Работы выполнены на высоком научно-методическом уровне и в срок.

OT MCM AH CCCP:

Мл. научный сотрудник

М. М. Кириченко 1905 1987r.

Мл. научный сотрудник <u> Г. А. Ляшко</u> 19.05 1987г.

От СКФ ВНИЦ "АИУС-агроресурсы":

Зав. отделом аэрокосмических и тематических изнсканий №4,к.э.н. Самсонов

-I987r. 19.05.

Гланий конструктор проекта
В. И. Коренец 1987r.

Главный конструктор проекта Е.В. Луценко 1987r.

Ho тогда семантическое ядро называлось позитивным информационным портретом класса, антияро – и негативные a информационным портретом.

3.2.3.3. Исследование предметной области путем исследования ее модели (задача 3)

3.2.3.3.1. Нелокальные нейроны и нейронные сети

В работе [19] автором предложены нелокальные нейроны и нелокальные нейронные сети прямого счета, основанные на теории информации. Эти сети, предложенные автором 15 лет очень сходны с популярным сегодня нейробайесовским подходом [20].

Экранная форма задания на формирование изображения нелокального нейрона приведена на рисунке 26, а пример самого нелокального нейрона группы специальностей: «06.02.00-Ветеринария» в СК-модели INF3 приведен на рисунке 27 (со всеми и 36 рецепторами), а фрагмент нелокальной нейронной сети в СК-модели INF3 на рисунке 28:

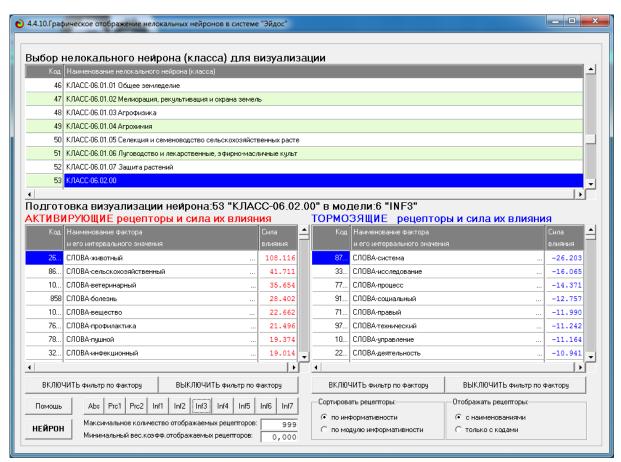
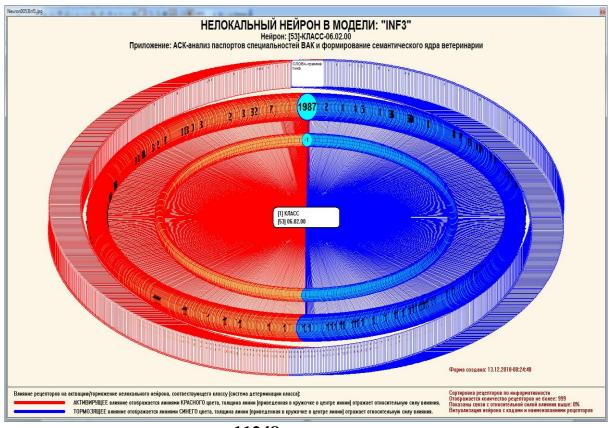
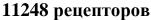
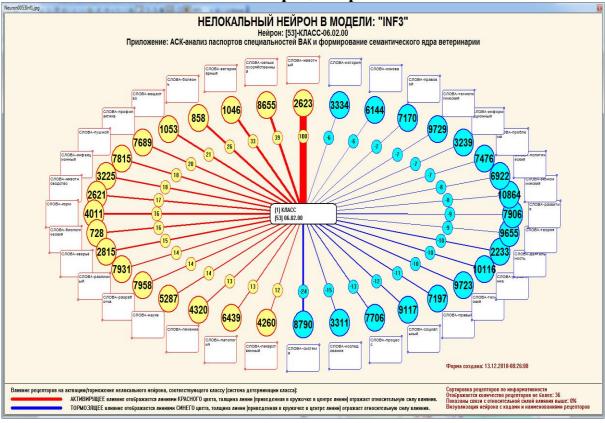


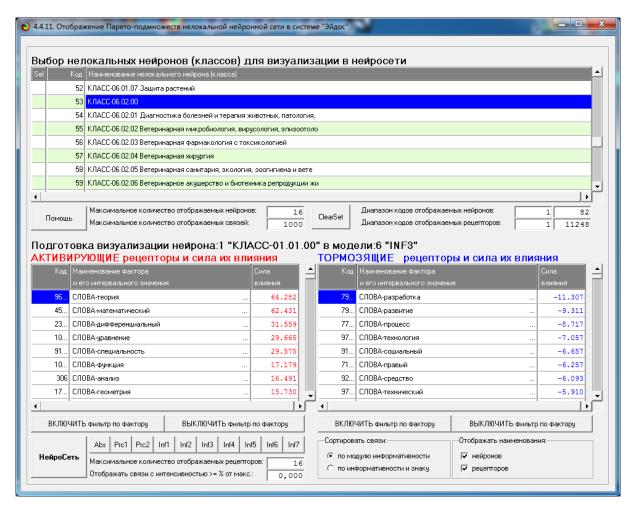
Рисунок 26. Экранная форма задания на формирование изображения нелокального нейрона







36 рецепторов Рисунок 27. Нелокальный нейрон



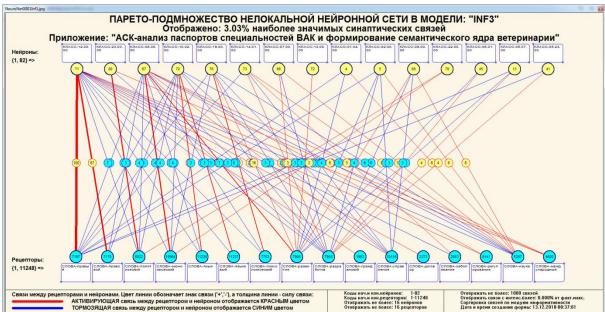


Рисунок 28. Экранная форма задания на визуализацию Парето-подмножества одного слоя нелокальной нейронной сети и сама визуализация (показано 3,03% наиболее значимых связей)

Отметим, что системно-когнитивные модели (СК-модели), созданные в системе «Эйдос», отражают смысл каждого слова, который в данном случае представляет собой количество информации (в битах), содержащееся в слове, о принадлежности текста с данным словом к тому или иному классу (рисунок 29):

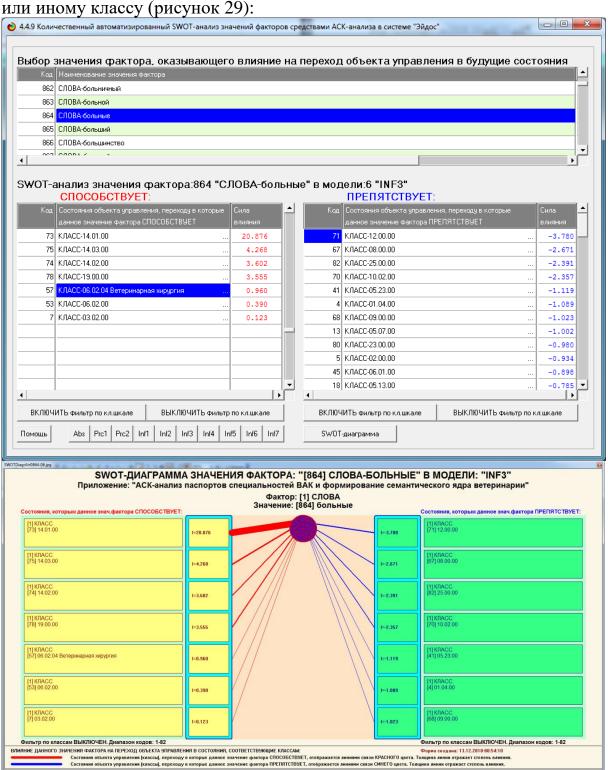


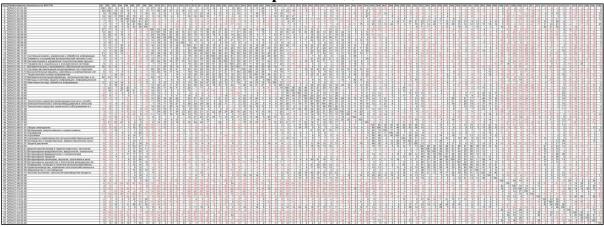
Рисунок 29. Семантический потенциал слова «больные» в системно-когнитивной модели INF3

Информация о сходстве и различии слов по смыслу также формируется в СК-моделях. Есть в системе «Эйдос» и выходные формы, отражающие эту информацию.

3.2.3.3.2. Когнитивные диаграммы и агломеративные дендрограммы классов

В режиме 4.2.2.1 рассчитывается матрица сходства классов, отражающая сходство-различие классов друг с другом в процентах (таблица 7).

Таблица 7 – Матрица сходства классов³⁸



В графической форме эта таблица отображается в форме когнитивных диаграмм (рисунок 30) и агломеративной дендрограммы с графиком изменения межкластерных расстояний (рисунок 31):

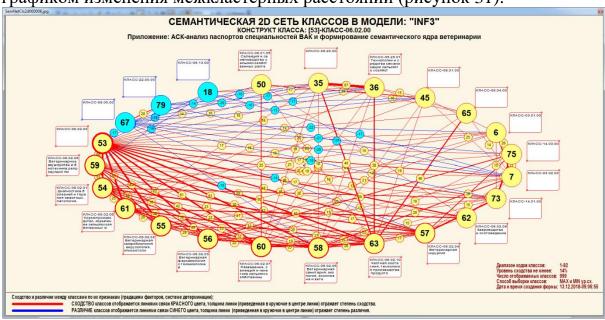
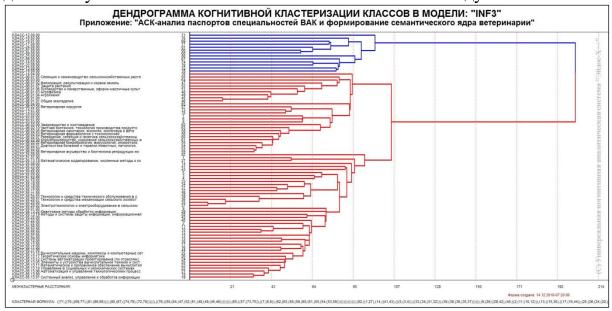


Рисунок 30. Когнитивная диаграмма, отражающая сходство-различие обобщенных образов классов друг с другом

 $^{^{38}}$ Изображение имеет разрешение 600 dpi и может быть просмотрено в масштабе 500%

Малозначащие связи могут отфильтрованы (не отображаться). В данном случае показаны связи с силой 14% и более по модулю.



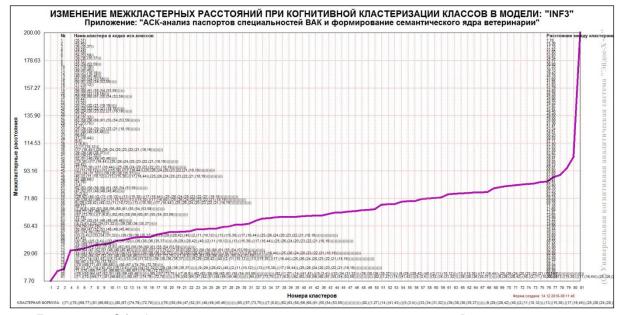


Рисунок 31. Агломеративная дендрограмма и график изменения межкластерных расстояний, отражающие сходство-различие обобщенных образов классов друг с другом [22, 18]

Из рисунков 30 и 31 видно, что классы образуют два кластера, являющиеся смысловыми полюсами конструкта (на рисунке 31 они выделены разными цветами: красным и синим). В кластеры объединены сходные тексты (паспорта специальностей), а на полюсах конструкта находятся противоположные по смыслу кластеры.

3.2.4. Выводы

3.2.4.1. Эффективность предложенного решения проблемы (задача 4)

Как показывает анализ результатов численного эксперимента предложенное и реализованное в системе «Эйдос» решение поставленных задач является вполне эффективным, что позволяет обоснованно утверждать, что цель работы достигнута.

3.2.4.2. Ограничения и недостатки предложенного решения проблемы и перспективы его развития путем их преодоления этих ограничений и недостатков (задача 5)

Однако, длительность решения задачи агломеративной когнитивной кластеризации для одной СК-модели при 82 классах и 11248 признаках составляет несколько часов. Это является неприемлемым. С другой стороны решение этой задачи в разумные сроки представляет большой научный и практический интерес. Решить эту проблему планируется в будущих работах.

3.2.4.3. Заключение

В результате проделанной работы, с помощью системы «Эйдос» были формированы обобщенные лингвистические образы классов по паспортам специальностей научных работников ВАК РФ и по группам специальностей (семантические ядра и антиядра), и, на основе этого, решены задачи идентификации текстов по направлениям науки, поддержки принятия решений и исследования моделируемой предметной области путем исследования ее модели.

Таким образом, созданы системно-когнитивные модели высокой достоверности, поставленные задачи решены, цель достигнута, проблема решена.

Данная работа демонстрирует, что математические модели (частные и интегральные критерии), методики численных расчетов (структуры данных и алгоритмы их обработки), экранные формы управления процессами, программный интерфейс ввода текстовых данных в систему «Эйдос» и повышения степени формализации исходных данных от вербализации до нормализованных баз данных (API), экранные формы текстовых и графических выходных форм по результатам решения задач глубокой атрибуции текстов, программная реализация математических моделей, методик численных расчетов, интерфейса и когнитивной графики в системе «Эйдос» являются адекватным средством для решения поставленной и решаемой в статье проблемы.

Необходимо системно-когнитивные отметить, ЧТО модели, разработанные в системе «Эйдос», могут быть применены для решения практических задач с применением той же системы «Эйдос», в которой они созданы, причем это применение возможно в адаптивном режиме, т.е. их можно совершенствовать в процессе эксплуатации, адаптировать к изменениям предметной области, локализовать или районировать для и т.п, т.д. Эти других регионов, И уникальные обеспечиваются тем, что система «Эйдос» представляет собой не только <u>среду</u> для эксплуатации интеллектуальных приложений, но и является <u>инструментом</u> их создания и адаптации.

По мнению автора АСК-анализ и система «Эйдос» представляют инновационный, т.е. доведенный ДО возможности практического применения, метод искусственного интеллекта может рассматриваться как универсальный инструмент решения всех тех задач в области ветеринарии (и других наук), для решения которых используется естественный интеллект. Причем ЭТО инструмент, многократно увеличивающий возможности естественного интеллекта, примерно также, как микроскоп и телескоп многократно увеличивает возможности естественного зрения, естественно только в том случае, если оно есть. Поэтому, конечно, этих задач огромное количество.

В качестве перспектив можно было бы отметить в частности решение следующих задач ветеринарии с применением автоматизированного системно-когнитивного анализа:

- поддержка принятия решений по выбору антибактериальных препаратов в зависимости от характера микробной флоры;
- поддержка принятия решений по определению дозы и пути введения препаратов группы пенициллина;
- поддержка принятия решений по определению дозы и пути введения цефалоспоринов;
- поддержка принятия решений по выбору антибактериальных препаратов с учетом основных токсических и аллергических реакций на антибактериальные препараты;
- исследование взаимодействия антибактериальных препаратов с другими препаратами при приеме внутрь и поддержка принятия решений по выбору антибактериальных препаратов с учетом результатов этих исследований.

Область ветеринарии, в которой перечисленные выше и другие задачи решаются с применением системно-когнитивного анализа,

программным инструментарием которого *в настоящее время* является система «Эйдос», предлагается назвать «*Когнитивной ветеринарией*» или шире «*Математической ветеринарией*», по аналогии с математической экономикой (08.00.13), Прикладной и математической лингвистикой (10.02.21), и т.д.

Эта идея находится в русле Указа Президента РФ от 7 июля 2011 г. N 899 "Об утверждении приоритетных направлений развития науки, технологий и техники в Российской Федерации и перечня критических технологий Российской Федерации", в котором под п.8 указаны Нано-, био-, информационные, когнитивные технологии 39.

Этим и другим применениям способствует и то, что система «Эйдос» является мультиязычной интеллектуальной on-line средой для обучения и научных исследований [2, 3]⁴⁰ и находится в полном открытом бесплатном доступе (причем с подробно комментированными актуальными исходными текстами: http://lc.kubagro.ru/_AIDOS-X.txt) на сайте автора по адресу: http://lc.kubagro.ru/aidos/_Aidos-X.htm.

Численные примеры решения задач ветеринарии с применением технологий искусственного интеллекта размещены как облачные Эйдосприложения под номерами: 100, 125, 126, 127, 128, 131, 133 и доступны всем желающим в режиме 1.3 системы «Эйдос». Базовое интеллектуальное приложение, являющееся основой данной работы, размещено в Эйдособлаке под номером 134.

Конечно, представленный в статье уровень исследования относится хотя и к развитому, но эмпирическому уровню, т.е. это просто наблюдаемые факты, эмпирические закономерности и в лучшем случае, подтверждения полученных результатов исследователями, может подняться до уровня эмпирического закона. Для перехода на теоретический уровень познания необходимо выдвинуть содержательной интерпретации полученных (которые может выдвинуть только специалист в области ветеринарии), объясняющие внутренние механизмы наблюдаемых закономерностей. Потом необходимо подтвердить, что эти научные гипотезы имеют силу, обнаружить прогностическую T.e. ПОЗВОЛЯЮТ неизвестные явления, и тогда эти гипотезы переходят в статус научной теории. Эта теория позволяют обобщить эмпирический закон до уровня научного закона [34].

http://lc.kubagro.ru/aidos/Presentation_Aidos-online.pdf

³⁹ Отметим, что все приведенные выше аргументы введения научного понятия: «когнитивная ветеринария» применимы и к другим направлениям науки, например: «когнитивная агрономия», «когнитивная экономика» и т.д.. Автор пытался развивать когнитивную математику [10] и когнитивную теорию управления [10], а также применять их в других областях науки и практики.

3.3. Агломеративная когнитивная кластеризация нозологических образов в ветеринарии

В статье на небольшом численном примере рассматривается сходство и различие нозологических образов в ветеринарии с применением нового метода агломеративной кластеризации, реализованного в автоматизированном системно-когнитивном анализе (АСК-анализ). Этот метод получил название: «Агломеративная когнитивная кластеризация». Этот метод отличается от известных традиционных тем, что: а) в нем параметры обобщенного образа кластера вычисляются не как средние от исходных объектов (классов) или их центр тяжести, а определяются с помощью той же самой базовой когнитивной операции АСК-анализа, которая применяется и для формирования обобщенных образов классов на основе примеров объектов и которая действительно корректно обеспечивает обобщение; б) в качестве критерия сходства используется не евклидово расстояние или его варианты, а интегральный критерий неметрической природы: «суммарное количество информации», применение которого теоретически корректно и дает хорошие результаты в неортонормированных пространствах, которые обычно и встречаются на практике; в) кластерный анализ проводится не на основе исходных переменных, матриц частот или матрицы сходства (различий), зависящих от единиц измерения по осям, а в когнитивном пространстве, в котором по всем осям (описательным шкалам) используется одна единица измерения: количество информации, и поэтому результаты кластеризации не зависят от исходных единиц измерения признаков объектов. Все это позволяет получить результаты кластеризации, понятные специалистам и поддающиеся содержательной интерпретации, хорошо согласующиеся с оценками экспертов, их опытом и интуитивными ожиданиями, что часто представляет собой проблему для классических методов кластеризации.

Раздел 3.6 основан на работе [6]. Ссылки на литературные источники в данном разделе даны по списку литературы работы [6].

В данной статье процесс диагностики и лечения животных рассматривается с позиций теории управления с применением искусственного интеллекта.

Состав классической системы управления с обратной связью приведен на рисунке 1:



Рисунок 6. Цикл управления в замкнутой системе управления

Конкретизируем стандартные универсальные термины теория управления для ветеринарии:

- объектом управления является животное;
- управляющая система это ветеринар;
- информация обратной связи о состоянии объекта управления это клиническая картина (непосредственно наблюдаемая симптоматика) и, возможно, данные лабораторных исследований;
- управляющие факторы это лечебные воздействия со стороны ветеринара на больное животное;
- целью управления является перевести животное из нежелательного состояния болезни в здоровое состояние.

Для принятия решения об управляющих воздействиях, т.е. для разработки *плана лечения*, необходимо предварительно поставить больному животному *нозологический диагноз* (от греч. nosos — болезнь и logos — слово, учение; букв.— учение о болезни), т.е. выработать заключение о сущности заболевания и состоянии животного, выраженное в нозологических терминах (установить болезненную форму или нозологическую единицу).

Задача нозологической диагностики может решаться в частности на основе знания симптоматики и по сути представляет собой задачу классификации объекта по его признакам (идентификация, распознавание). Решение этой задачи в ветеринарии (как и задач: поддержки управляющих решений и исследования моделируемой предметной области путем

исследования ее модели) подробнейшим образом рассмотрено в работе [1]. Для решения этих задач применен автоматизированный системно-когнитивный анализ (АСК-анализ) и его программный инструментарий Универсальная автоматизированная система «Эйдос» [2, 3]⁴¹.

В данной работе рассматривается и решается **проблема**, возникающая при нозологической диагностике (классификации), состоящая в том, что различные нозологические единицы в ряде случаев имеют сходную симптоматику и являются сложно различимыми из-за чего возникает опасность постановки неверного диагноза и, соответственно, неправильного лечения ⁴².

Предлагается идея решения проблемы состоящая в том, что ветеринар должен иметь информацию о степени сходства-различия нозологических образов по симптоматике должен учитывать ее при интерпретации выходных форм по результатам диагностики. Если говорить более конкретно, то ветеринар понимая этиологию различных заболеваний может видеть глубокую внутреннюю взаимосвязь между ними и иногда может увидеть причину симптоматически более ярко выраженного заболевания в другом заболевании, симптоматика которого выражена менее четко.

Отметим, что первоначально под нозологией понимали раздел патологии, включающий общее учение о болезни (общая нозология), а также изучающий этиологию, патогенез клинические и патологоанатомические проявления отдельных болезней (частная нозология), а также *классификацию* и номенклатуру болезней ⁴³.

В качестве примера рассмотрим результат нозологической диагностики, приведенный на рисунке 21 из работы [1], отображающий сходство клинической картины конкретного пациента («Мурка») с обобщенными нозологическими образами (рисунок 2)/

Возникает вопрос, от какого заболевания лечить Мурку, от 1-го в этом списке, или 2 или может быть даже 3-го.

Для ответа на этот вопрос применим кластерный анализ нозологических образов.

Кластерный анализ (англ. *Data clustering*) — это задача разбиения заданной выборки *объектов* (ситуаций) на подмножества, называемые кластерами, так, чтобы каждый кластер состоял из схожих объектов, а объекты разных кластеров существенно отличались.

-

⁴¹ http://lc.kubagro.ru/aidos/Presentation Aidos-online.pdf

⁴² Известно, что лечение не от той болезни иногда может быть привести к более плачевным результатам, чем отсутствие лечения. Это может быть связано как с весьма нежелательными побочными эффектами применяемых методов, так и в очень большой степени просто с потерей ценнейшего времени.

⁴³ В этом понимании нозология не имеет чётких границ с патологией.

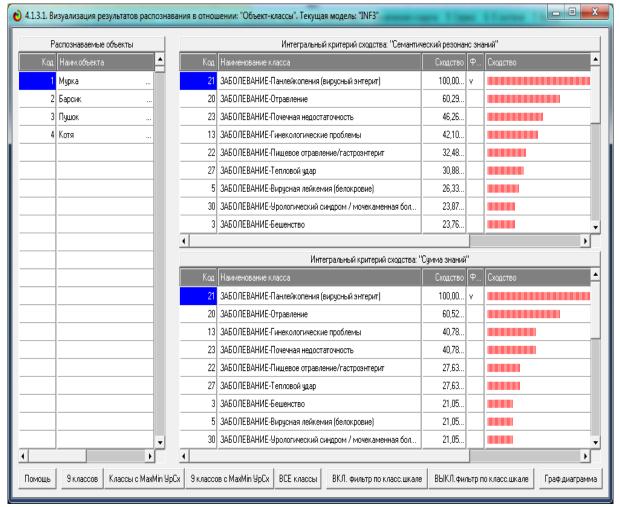


Рисунок 7. Экранная форма с результатами нозологической диагностики пациента «Мурка», (симптоматика из обучающей выборки)

Однако мы не будем применять классические методы кластерного анализа и реализующий их программный инструментарий. Причиной этого является то, что эти методы имеют ряд довольно существенных на взгляд автора недостатков. Эти недостатки подробно проанализированы в работе [4], в ней же предложено и их решение [4, 5]. Поэтому в данной статье мы не будем их подробно рассматривать. Отметим лишь, что многие из этих недостатков, являющихся проблемами классических методов кластерного анализа, преодолены в АСК-анализе и интеллектуальной системе «Эйдос» (таблица 1).

Таблица 8 – ПРОБЛЕМЫ КЛАССИЧЕСКИХ МЕТОДОВ КЛАСТЕРИЗАЦИИ И ИХ РЕШЕНИЕ, ПРЕДЛАГАЕМЫЕ В АСК-АНАЛИЗЕ И СИСТЕМЕ «ЭЙДОС»

No	Формулировка проблемы	Решение, пПредлагаемое в АСК-анализе		
01-	кластерного анализа	и системе «Эйдос»		

1.	<u>Проблема 1.1</u> выбора метрики, корректной	Предлагается применить неметрический		
1.	для неортонормированных пространств.	интегральный критерий, представляющий		
	олл неортонормированных пространств.	собой суммарное количество информации в		
		системе признаков о принадлежности объекта		
		к классу («информационное расстояние»),		
		никак не основанный на предположении об		
		ортонормированности пространства и		
		корректно работающий в		
2	H (10	неортонормированных пространствах		
2.	<u>Проблема 1.2</u> ортонормирования пространства.	Поддерживается		
3.	Проблема 2.1 сопоставимой обработки	Объекты формально описываются в виде		
	описаний объектов, описанных признаками	векторов, затем рассчитывается матрица		
	различной природы, измеряемыми в различных	абсолютных частот и на ее основе – матрица		
	единицах измерения (проблема размерностей).	знаний, с использованием которой все		
	(признаки измеряются в одних единицах		
		измерения: единицах измерения количества		
		данных, информации и знаний – битах,		
		байтах, и т.д.		
	Проблема 2.2 формализации описаний	Числовые шкалы преобразуются в		
	объектов, имеющих как количественные, так	интервальные значения, после чего градации		
	и качественные признаки.	всех типов шкал обрабатываются		
	и кичественные призники.	единообразно (см.п.3)		
4.	<u>Проблема 3.1</u> доказательства гипотезы о	Нет необходимости, т.к. предлагаемые		
4.				
	нормальности исходных данных.	частные и интегральные критерии не		
		предполагают нормальности исходных		
-	П (22	данных		
5.	<u>Проблема 3.2</u> нормализации исходных данных.	Реализованы режимы ремонта или		
	H ()	взвешивания исходных данных.		
6.	<u>Проблема 3.3</u> применения непараметрических	Предлагаемые методы являются		
	методов кластеризации, корректно	непараметрическими и корректно работают с		
	работающих с ненормализованными данными.	ненормализованными данными		
7.	<u>Проблема 4</u> разработки такого метода	Предлагаемый метод включает фильтр		
	кластерного анализа, математическая модель	подавления шума на уровне формирования		
	и алгоритм и которого органично включали	матрицы абсолютных частот и самой		
	бы фильтр, подавляющий шум в исходных	математической форме интегрального		
	данных, в результате чего данный метод	критерия. Кроме того, реализованы режимы		
	кластеризации корректно работал бы при	удаления или корректной обработки		
	наличии шума в исходных данных.	артефактов, выбросов (нетипичных объектов)		
		и малопредставленных данных, по которым		
		нет достаточной статистики в исходных		
		данных		
8.	Проблема 5 разработки метода кластерного	Поддерживается исключение выбросов и		
	анализа, математическая модель и алгоритм	артефактов из дендрограмм, либо их		
	и которого обеспечивали бы <u>выявление</u>	отображение специальным для них образом.		
	«выбросов» (артефактов) в исходных данных			
	и позволяли либо вообще не показывать их в			
	дендрограммах, либо показывать, но так,			
	чтобы было наглядно видно, что это			
	артефакты.			

В статье [4] были отмечены и недостатки самого предложенного в ней решения, а также указаны перспективы его развития. Эти недостатки связаны прежде всего с программной реализацией DOS-версии самой системы «Эйдос», которая была текущей на момент написания работы [4]. В текущей версии системы «Эйдос» от 21.04.2018 все эти ограничения

DOS-версии преодолены и реализованы многие из указанных в [4] перспектив. В частности реализован режим агломеративной когнитивной кластеризации (режим 4.2.2.3). Вместо описания данного режима приведем его Help (рисунок 3):

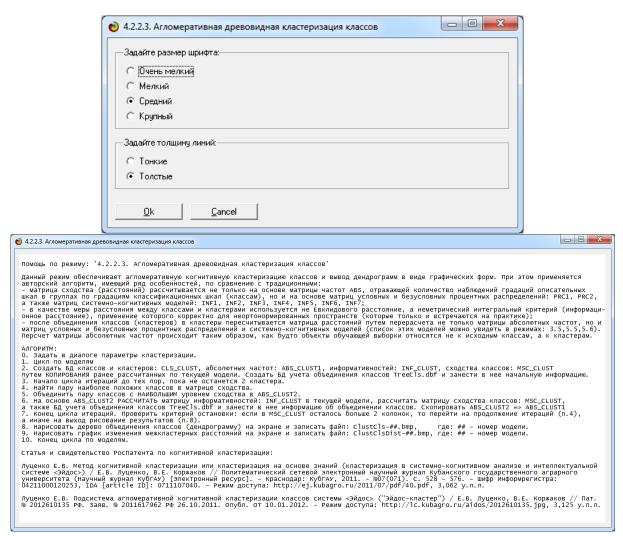
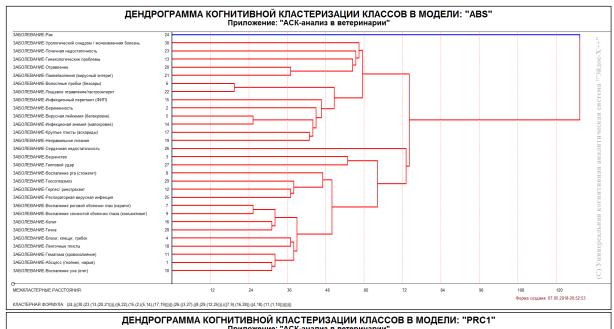
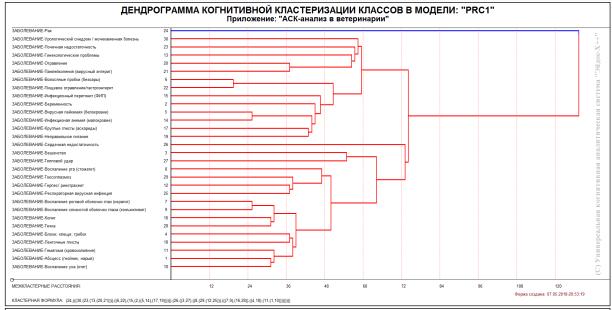
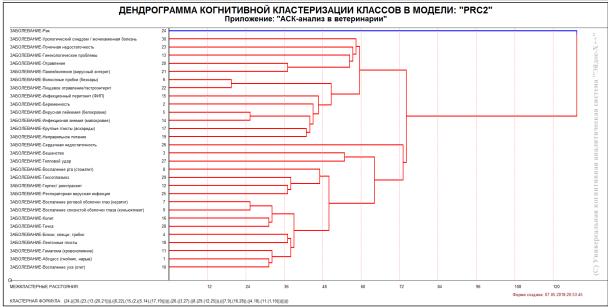


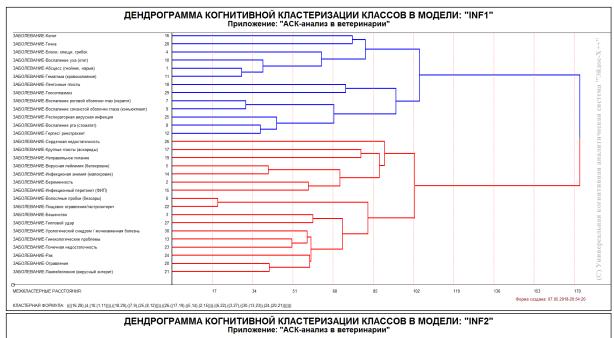
Рисунок 3. окно настроек визуализации дендрограмм и Help режима агломеративной когнитивной кластеризации

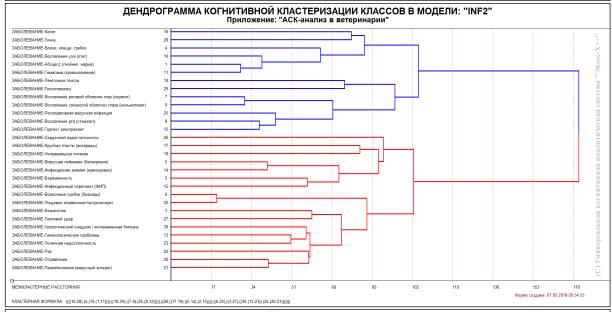
На рисунках 4 приведены агломеративные дендрограммы нозологических образов, построенные на основе моделей, созданных и исследованных в работе [1].

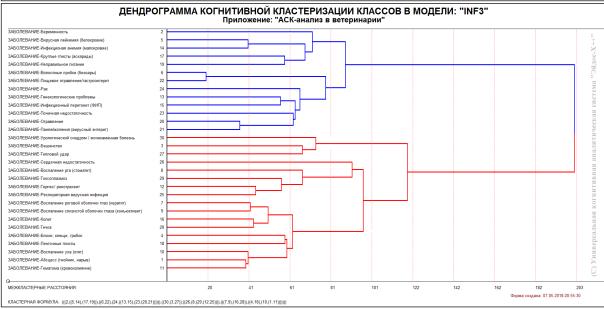


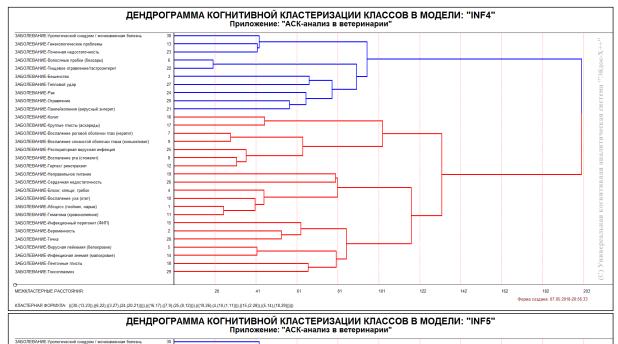


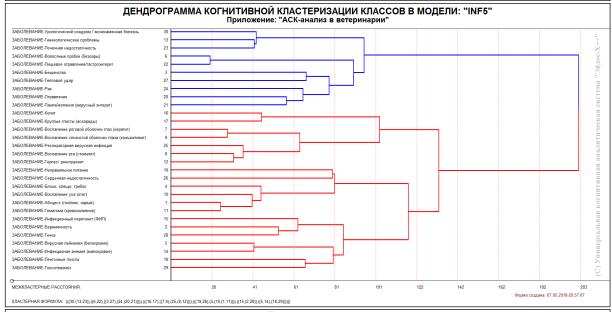


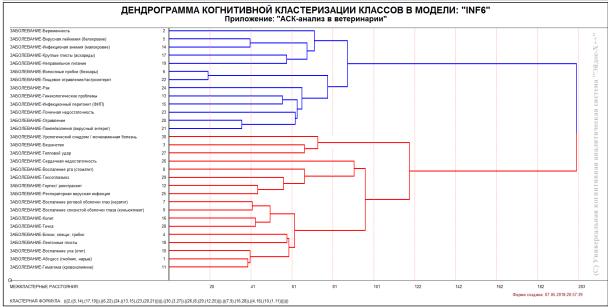












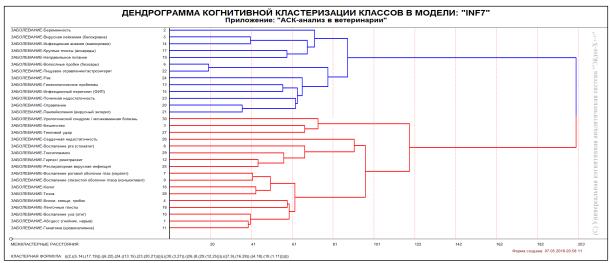
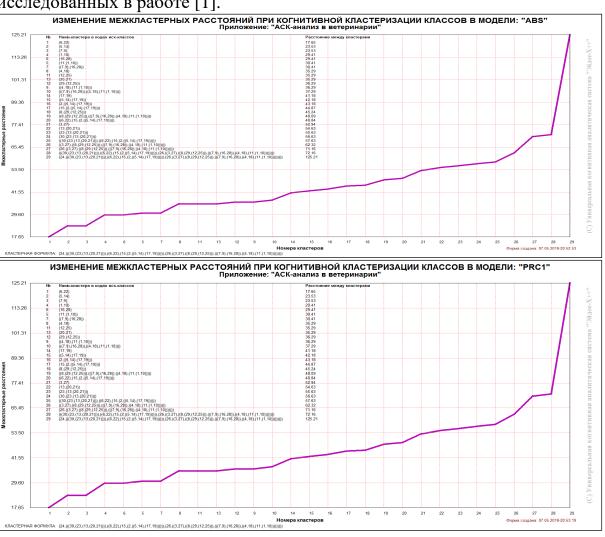
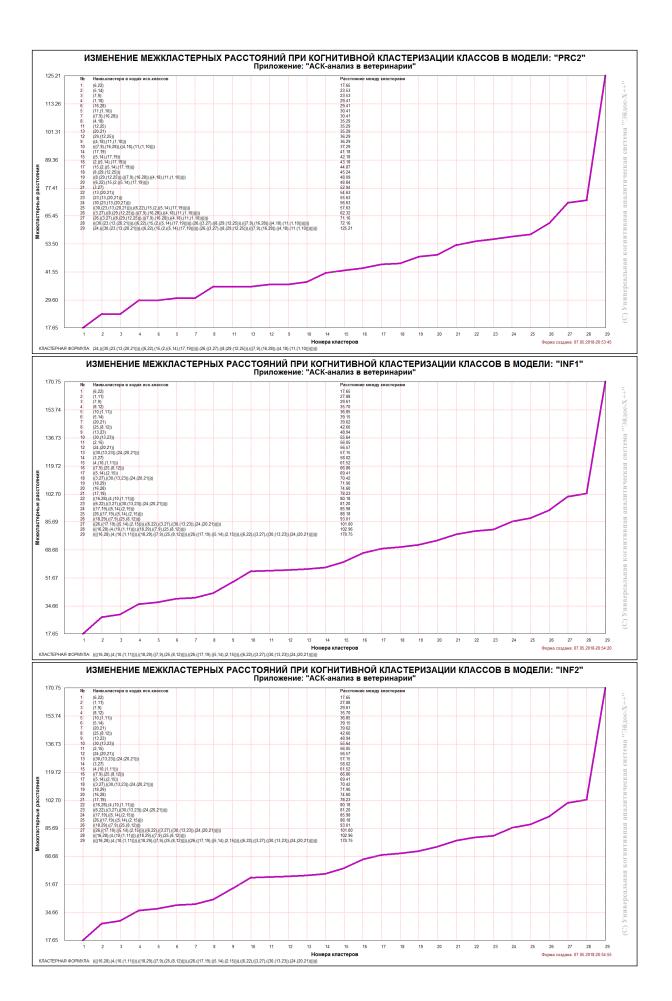
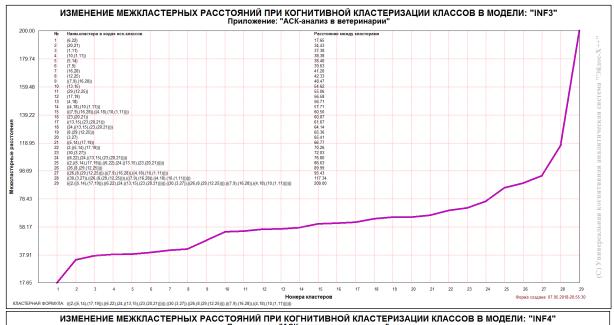


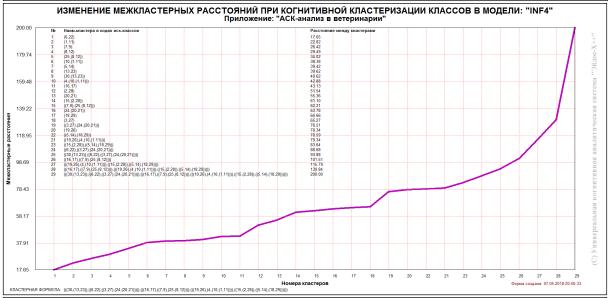
Рисунок 4. Дендрограммы нозологических образов, построенные на основе моделей, созданных и исследованных в работе [1]

На рисунках 5 приведены графики межкластерных расстояний нозологических образов, построенные на основе моделей, созданных и исследованных в работе [1].











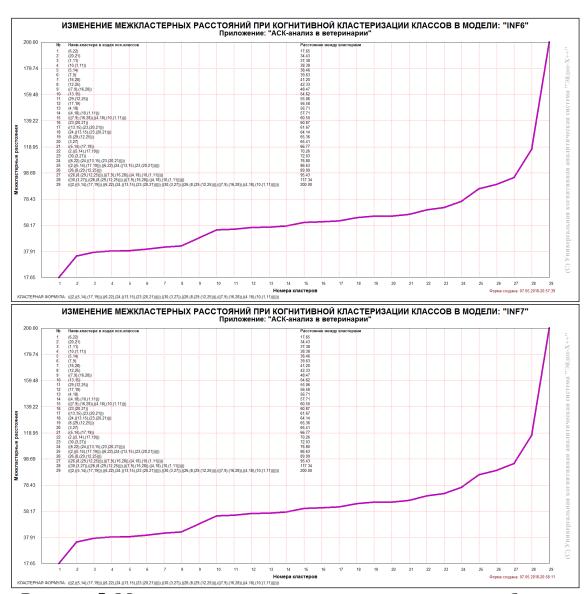


Рисунок 5. Межкластерные расстояния нозологических образов, построенные на основе моделей, созданных и исследованных в работе [1]

Отметим, что в традиционных формах агломеративного кластерного анализа есть *частичный* аналог дендрограммы в модели ABS, а дендрограммы, аналогичные полученным в моделях PRC1, PRC2 и в системно-когнитивных моделях: INF1, INF2, INF3, INF4, INF5, INF6 и INF7 вообще отсутствуют.

Результаты кластерно-конструктивного анализа классов отображаются также в форме когнитивных диаграмм, пример одной из которых приведен на рисунке 6:

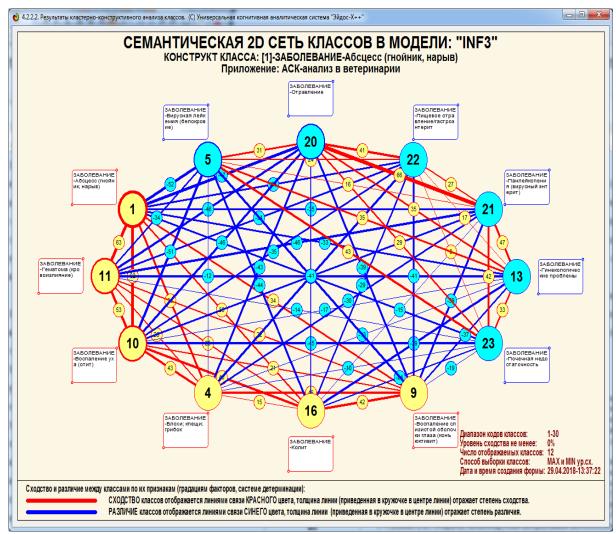


Рисунок 8. Пример когнитивной диаграммы с результатами кластерно-конструктивного анализа классов

По результатам статьи можно сделать обоснованный **выво**д о том, что метод «Агломеративной когнитивной кластеризации» является адекватным средством решения поставленной проблемы, т.к. позволяет достоверно оценить сходство-различие нозологических образов в ветеринарии по их симптоматике. Это позволяет ветеринару более обоснованно применять решение о назначении лечения с учетом не только результатов диагностики, но и классификации самих нозологических образов и их этиологии.

В качестве недостатка предлагаемого метода можно указать на более значительные затраты времени и вычислительных ресурсов по сравнению с классическими традиционными методами кластеризации.

Как **перспективу** развития метода отметим: повышение размерности дендрограмм и разработку режима агломеративной когнитивной кластеризации признаков.

Материалы данной статьи могут быть использованы в учебном процессе при преподавании дисциплин: «Интеллектуальные

информационные системы», «представление знаний в информационных системах», «Современные технологии в образовании (магистратура)», «Управление знаниями (магистратура)», «Введение в искусственный интеллект», «Интеллектуальные и нейросетевые технологии в образовании (магистратура)», «Основы искусственного интеллекта» и других.

Планируется продолжить исследования и разработки, представленные в данной статье, по следующим направлениям:

- 1. Агломеративная когнитивная кластеризация симптомов и синдромов в ветеринарии.
- 2. Разработка в среде интеллектуальной системы Эйдос-X++ ветеринарных диагностических тестов и проведение научных исследований на основе эмпирических данных (данных наблюдений и экспериментов).
- 3. Интеграция в среде интеллектуальной системы Эйдос-X++ без программирования нескольких диагностических тестов в батарею тестов или в супертест с одной общей моделью.

Этим и другим применениям способствует и то, что система «Эйдос» является мультиязычной интеллектуальной on-line средой для обучения и научных исследований [3, 4]⁴⁴ и находится в полном открытом бесплатном доступе (причем с подробно комментированными открытыми исходными текстами: http://lc.kubagro.ru/AIDOS-X.txt) на сайте автора по адресу: http://lc.kubagro.ru/aidos/_Aidos-X.htm.

3.4. Агломеративная когнитивная кластеризация симптомов и синдромов в ветеринарии

В статье на небольшом численном примере рассматривается сходство и различие симптомов и синдромов по их диагностическому смыслу, т.е. по той информации, которую они содержат о принадлежности состояний животных к различным нозологическим образам. Эта задача решается для ветеринарии с применением нового агломеративной когнитивной кластеризации, реализованного автоматизированном системно-когнитивном анализе (АСК-анализ). Этот метод кластеризации отличается от известных традиционных тем, что: а) в нем параметры обобщенного образа кластера вычисляются не как средние от исходных объектов (симптомов) или их центр тяжести, а определяются с помощью той же самой базовой когнитивной операции АСК-анализа, которая применяется и для формирования обобщенных образов классов на основе примеров объектов и которая действительно корректно обеспечивает обобщение; б) в качестве критерия сходства используется не евклидово расстояние или его варианты, а интегральный критерий неметрической природы: «суммарное количество информации», применение которого теоретически корректно и дает хорошие результаты в неортонормированных пространствах, которые как правило и встречаются на практике; в) кластерный анализ проводится не на основе исходных переменных, матрии частот или матрицы

.

⁴⁴ http://lc.kubagro.ru/aidos/Presentation Aidos-online.pdf

сходства (различий), зависящих от единиц измерения по осям (измерительным шкалам), а в когнитивном пространстве, в котором по всем осям используется одна единица измерения: количество информации, и поэтому результаты кластеризации не зависят от исходных единиц измерения признаков объектов. Все это позволяет получить результаты кластеризации, понятные специалистам и поддающиеся содержательной интерпретации, хорошо согласующиеся с оценками экспертов, их опытом и интуитивными ожиданиями, что часто представляет собой проблему для классических методов кластеризации.

Раздел 3.7 основан на работе [7]. Ссылки на литературные источники в данном разделе даны по списку литературы работы [7].

Задача нозологической диагностики в ветеринарии может решаться на основе знания $симптоматики^{45}$ и по сути представляет собой задачу классификации состояния объекта по его признакам (идентификация, распознавание).

Симптомы можно классифицировать на

- *детерминистские*, взаимно-однозначно связанные с определенными заболеваниями, т.е. встречающимися только при них и при них встречающихся всегда (эти симптомы наиболее ценны для диагностики);
- *статистические*, которые чаще наблюдаются при одних заболеваниях, а при других реже (эти симптомы менее ценны для диагностики, чем детерминистские);
- *бесполезные*, которые всегда встречаются при всех заболеваниях, но отсутствуют у здоровых животных (наличие этих симптомов позволяет установить лишь сам факт заболевания, но не позволяет установить какое оно именно).

Большинство симптомов относится к статистическим, а детерминистские и бесполезные симптомы представляют собой два полюса к которым в различной степени приближаются наиболее ценные и наименее ценные статистические симптомы.

Существование статистических симптомов приводит к ситуации, при которой один и тот же симптом несет информацию о различных заболеваниях, но само это количество информации разное, для разных заболеваний, т.е. об одних заболеваниях статистический симптом несет больше информации, а о других меньше. Таким образом, различные симптомы могут иметь сходный диагностический смысл, т.е. нести информацию о принадлежности состояния животного к одним и тем же

⁴⁵ Симптом (греч. symptōma совпадение, случай, признак) признак болезни, обнаруживаемый с помощью клинических методов исследования и используемый для диагностики и (или) прогноза заболевания https://dic.academic.ru/dic.nsf/enc medicine/28346/симптом

нозологическим образам. Группа симптомов, сходных по диагностическому смыслу образует $cundpom^{47}$.

В этой связи возникает задача количественного определения сходства-различия симптомов по их диагностическому смыслу, т.е. задача выявления синдромов.

Решение данной задачи представляет интерес как для ученых, разрабатывающих и совершенствующих методы и методики количественной диагностики в ветеринарии, так и для врачей-ветеринаров, применяющих эти методы диагностики на практике.

Для решения поставленной задачи может быть применен метод кластерного анализа.

Кластерный анализ (англ. *Data clustering*) — это задача разбиения заданной выборки *объектов* (ситуаций) на подмножества, называемые кластерами, так, чтобы каждый кластер состоял из схожих объектов, а объекты разных кластеров существенно отличались.

Однако мы не будем применять классические методы кластерного анализа и реализующий их программный инструментарий. Причиной этого является то, что эти методы имеют ряд довольно существенных (на взгляд автора) недостатков. Эти недостатки подробно проанализированы в работе [1], в ней же предложено и их решение [1, 2]. Поэтому в данной статье мы не будем их подробно рассматривать. Отметим лишь, что многие из этих недостатков, являющихся проблемами классических методов кластерного анализа (таблица 1 предыдущего раздела), преодолены в новом инновационном (т.е. доведенным до практического применения) методе искусственного интеллекта: автоматизированном системно-когнитивном анализе (АСК-анализ) и его программном инструментарии: Универсальной автоматизированной системе «Эйдос» [3, 4]⁴⁸.

В работе [1] были отмечены также и недостатки самого предложенного в ней решения, т.е. указаны перспективы его развития. Эти недостатки связаны прежде всего с программной реализацией DOS-версии самой системы «Эйдос», которая была текущей на момент написания работы [1]. В новой версии системы «Эйдос» от 05.06.2018 все эти ограничения DOS-версии преодолены и реализованы многие из указанных в [1] перспектив. В частности реализован режим агломеративной когнитивной кластеризации классов (режим 4.2.2.3) и признаков (режим 4.3.2.3).

Применим режим агломеративной когнитивной кластеризации признаков для решения поставленной в данной работе задачи. При этом мы используем интеллектуальное приложение системы «Эйдос»,

48 http://lc.kubagro.ru/aidos/Presentation_Aidos-online.pdf

⁴⁶ от греч. nosos — болезнь и logos — слово, учение; букв. — учение о болезни.

⁴⁷ Синдром (syndromom, греч. syndromē стечение, скопление) устойчивая совокупность ряда симптомов с единым патогенезом: https://dic.academic.ru/dic.nsf/enc_medicine/28878/cuндром

описанное в статьях [5, 6]. Отметим также, что данное приложение размещено в облаке на ftp-сервере системы «Эйдос» и является современным развитием ранее проведенной работы [7].

Вместо описания режима 4.3.2.2 приведем его Help (рисунок 1):

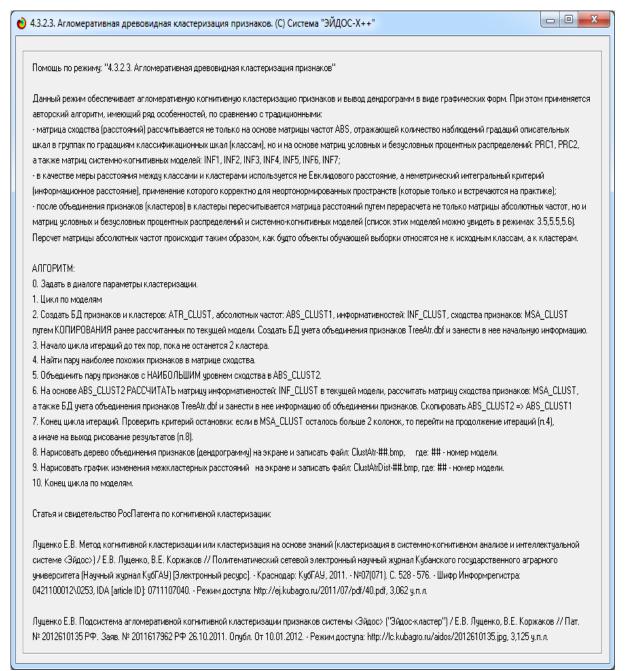
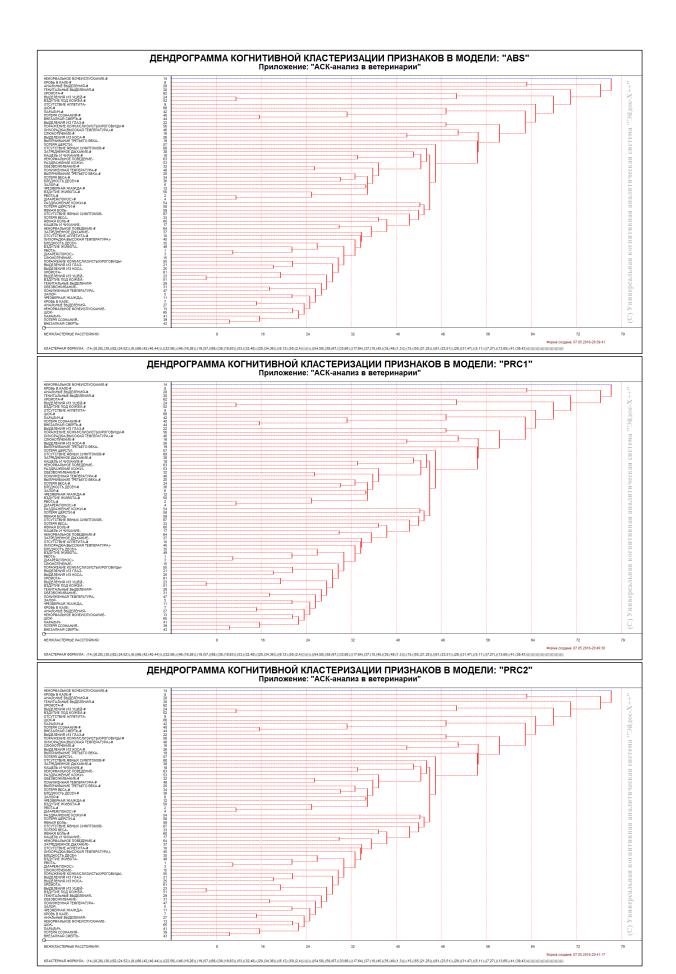
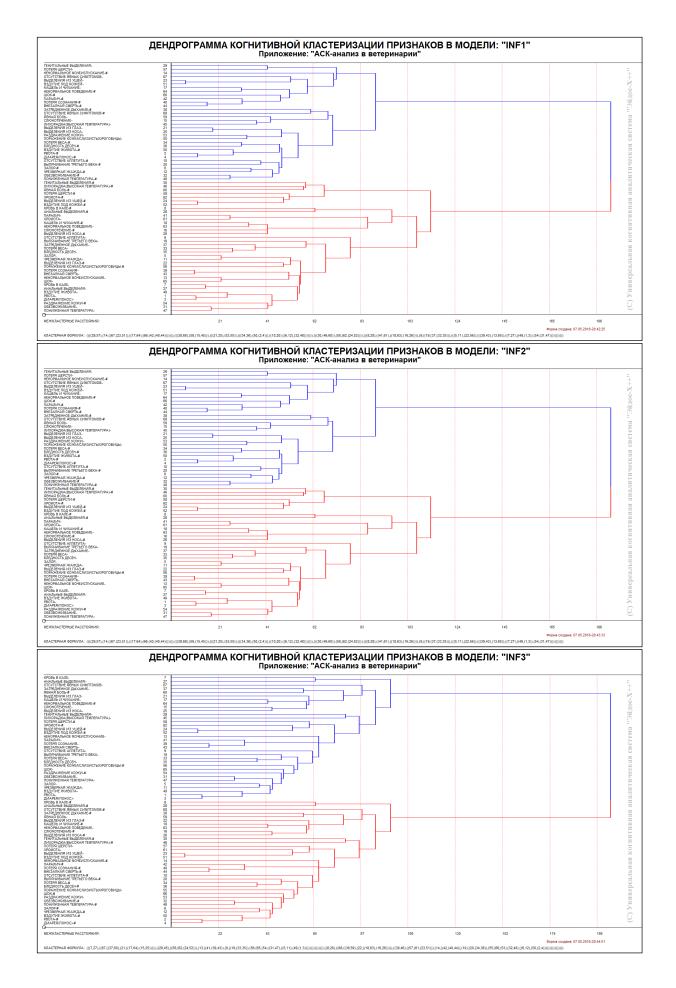
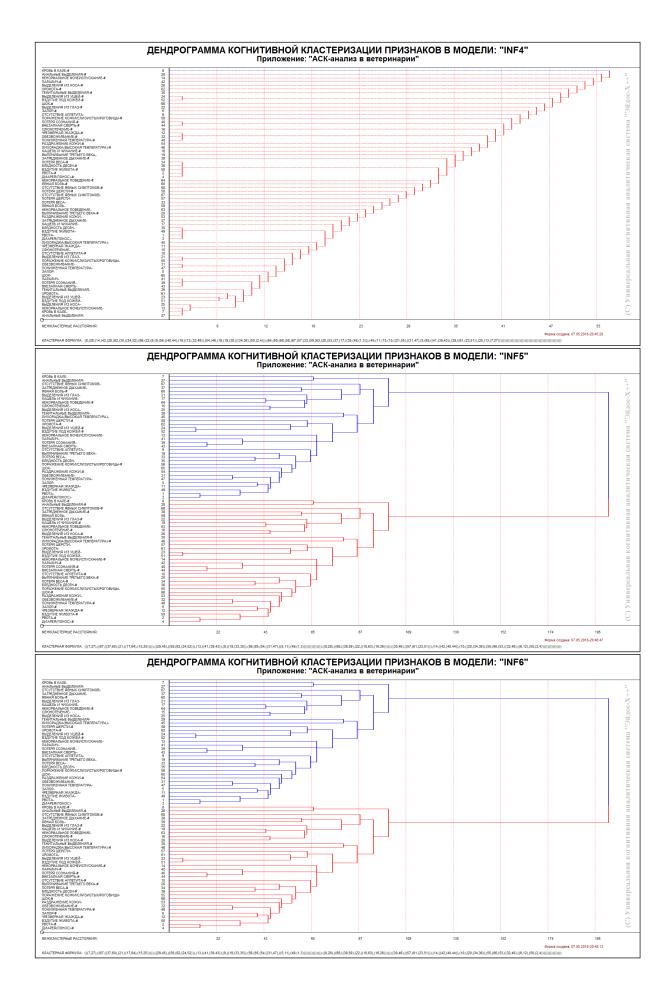


Рисунок 1. Help режима агломеративной когнитивной кластеризации признаков

На рисунках 2 приведены агломеративные дендрограммы симптомов и синдромов. При этом под синдромом, как мы уже говорили выше, мы будем понимать несколько симптомов, объединенных в группу по сходству их диагностического смысла.







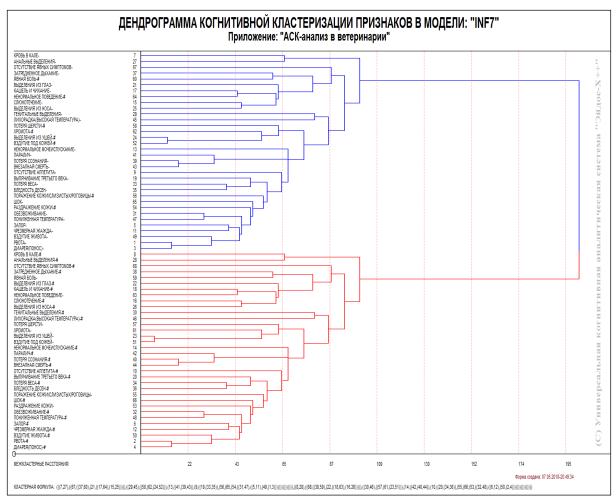
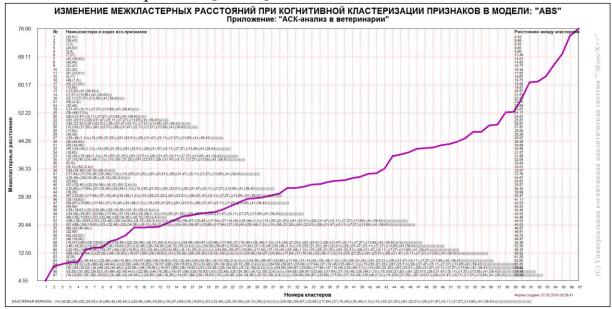
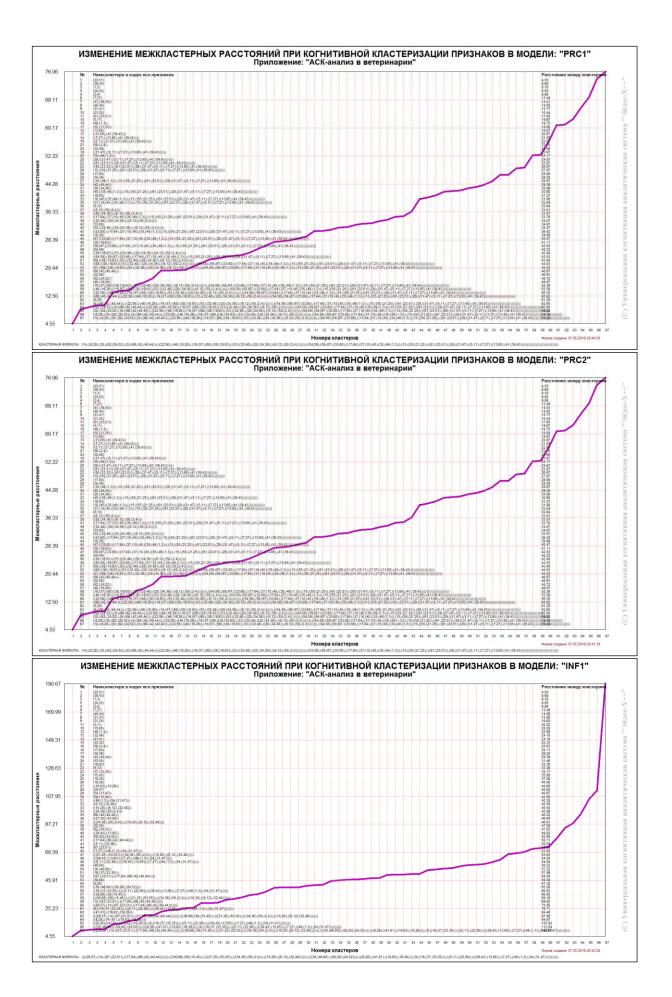
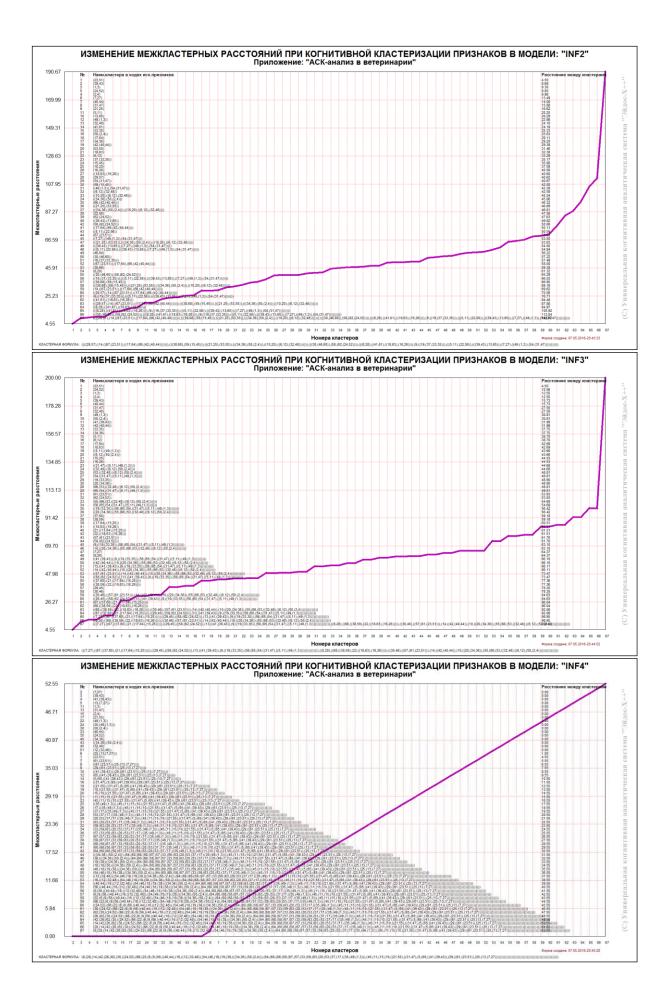


Рисунок 29. Дендрограммы симптомов и синдромов, построенные на основе моделей, созданных и исследованных в работах [5, 6, 7]

На рисунках 3 приведены графики межкластерных расстояний симптомов и синдромов, построенные на основе моделей, созданных и исследованных в работах [5, 6, 7].







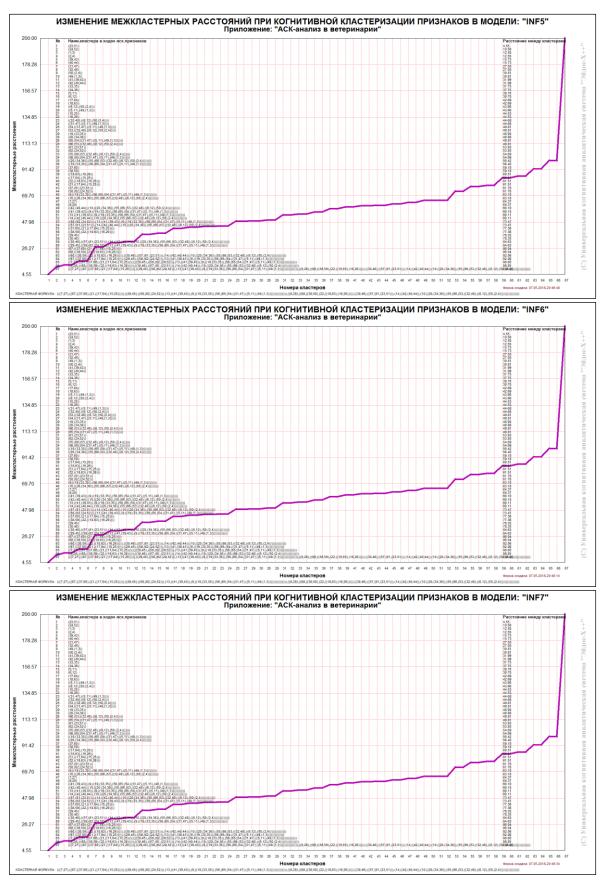


Рисунок 3. Межкластерные расстояния симптомов и синдромов, построенные на основе моделей, созданных и исследованных в работах [5, 6, 7]

Отметим, что в традиционных формах агломеративного кластерного анализа есть *частичный* аналог дендрограммы в модели ABS, а дендрограммы, аналогичные полученным в моделях PRC1, PRC2 и в системно-когнитивных моделях: INF1, INF2, INF3, INF4, INF5, INF6 и INF7 вообще отсутствуют.

Результаты кластерно-конструктивного анализа классов отображаются также в форме когнитивных диаграмм, пример одной из которых приведен на рисунке 4:

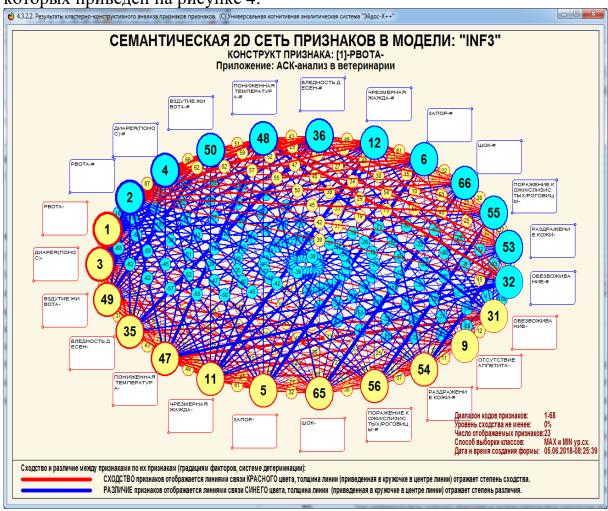


Рисунок 410. Форма управления и когнитивная диаграмма с результатами кластерно-конструктивного анализа симптомов

По результатам статьи можно сделать обоснованный **вывод** о том, что метод «Агломеративной когнитивной кластеризации» является адекватным средством решения поставленной задачи, т.к. позволяет достоверно оценить сходство-различие симптомов и синдромов в ветеринарии по их диагностическому смыслу. Это позволяет ветеринару более обоснованно применять решение о назначении лечения с учетом не только результатов диагностики, но и классификации самих симптомов и синдромов.

В качестве **недостатка** предлагаемого метода можно указать на более значительные затраты времени и вычислительных ресурсов по сравнению с классическими традиционными методами кластеризации.

Как **перспективу** развития метода отметим: повышение размерности дендрограмм и разработку других видов дендрограмм, например, круговых.

Материалы данной статьи могут быть использованы в учебном процессе при преподавании дисциплин: «Интеллектуальные информационные системы», «представление знаний в информационных системах», «Современные технологии в образовании (магистратура)», «Управление знаниями (магистратура)», «Введение в искусственный интеллект», «Интеллектуальные и нейросетевые технологии в образовании (магистратура)», «Основы искусственного интеллекта» и других.

Планируется продолжить исследования и разработки, представленные в данной статье, по следующим направлениям:

- 1. Агломеративная когнитивная кластеризация нозологических образов, симптомов и синдромов в ветеринарии на конкретных примерах, относящихся к другим животным и группам заболеваний.
- 2. Разработка в среде интеллектуальной системы Эйдос-X++ ветеринарных диагностических тестов и проведение научных исследований на основе эмпирических данных (данных наблюдений и экспериментов).
- 3. Интеграция в среде интеллектуальной системы Эйдос-X++ без программирования нескольких диагностических тестов в батарею тестов или в супертест с одной общей моделью.

Этим и другим применениям способствует и то, что система «Эйдос» является мультиязычной интеллектуальной on-line средой для обучения и научных исследований $[3, 4]^{49}$ и находится в полном открытом бесплатном доступе (причем с подробно комментированными открытыми исходными текстами: http://lc.kubagro.ru/_AIDOS-X.txt) на сайте автора по адресу: http://lc.kubagro.ru/aidos/_Aidos-X.htm.

4. ОБСУЖДЕНИЕ

Полученные результаты можно оценить как успешно решающие сформулированную в работе проблему и обеспечивающие достижение поставленной в работе цели. Эти результаты получены путем применения Автоматизированного системно-когнитивного анализа (АСК-анализ) и его программного инструментария — интеллектуальной системы «Эйдос».

Достижением данной работы является:

-

^{49 &}lt;u>http://lc.kubagro.ru/aidos/Presentation_Aidos-online.pdf</u>

- 1. Возможность построения системно-когнитивных моделей предметной области на основе исходных данных, в т.ч. содержащих лингвистические переменные.
- 2. Возможность применения системно-когнитивных моделей для решения задач прогнозирования, принятия решений и исследования моделируемой предметной области.

В качестве перспективы продолжения исследований можно было бы рекомендовать существенно увеличить количество и разнообразие решаемых задач и объем исходных данных, количество исследуемых факторов, а также количество классификационных шкал и их градаций (классов) для описания будущих состояний объекта моделирования.

Перспективность и ценность результатов подобных исследований и разработок для теории и практики не вызывает особых сомнений, что подтверждается работами авторов в этой области [1-14].

У желающих есть все возможности для изучения данной работы и для дальнейших исследований с применением АСК-анализа и системы «Эйдос» на своем компьютере.

Для этого надо скачать систему с сайта разработчика по ссылке на странице: http://lc.kubagro.ru/aidos/_Aidos-X.htm, а затем в диспетчере приложений (режим 1.3) установить соответствующие интеллектуальные облачные Эйдос-приложения. По различным аспектам применения данной технологии есть большое количество видео-занятий (около 300), с которыми можно ознакомиться по ссылкам, приведенным на странице: http://lc.kubagro.ru/aidos/How_to_make_your_own_cloud_Eidos-application.pdf.

5. ВЫВОДЫ

Главный вывод, который можно обоснованно сделать на основе постановки и решения рассмотренных задач в области ветеринарии состоит в том, что Автоматизированный системно-когнитивный анализ (АСК-анализ) и его программный инструментарий, в качестве которого в настоящее время выступает интеллектуальная система «Эйдос», являются инструментом познания, многократно увеличивающим его возможности, который можно успешно применять для решения разнообразных задач идентификации, диагностики, прогнозирования, принятия решений и исследования объекта моделирования путем исследования его модели в области ветеринарии.

Предлагается ввести новое научное понятие «Когнитивная ветеринария», под которой понимается ветеринария, широко использующая Автоматизированный системно-когнитивный анализ и другие методы искусственного интеллекта.

Содержание монографии соответствует образовательным программам по направлениям подготовки: 09.03.02 Информационные системы и технологии бакалавриата и 09.04.02 Информационные системы

и технологии магистратуры. Монография может быть использована в качестве дополнительной литературы при изучении дисциплин, по этим образовательным программам, связанным с интеллектуальными технологиями.

СПИСОК ЛИТЕРАТУРЫ

- 1. Луценко, Е. В. Когнитивная ветеринария ветеринария цифрового общества: дефиниция базовых понятий / Е. В. Луценко, Е. К. Печурина, А. Э. Сергеев // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. − 2019. − № 152. − С. 141-199. − EDN LQJMQH.
- 2. Луценко, Е. В. Формирование семантического ядра ветеринарии путем автоматизированного системно-когнитивного анализа паспортов научных специальностей ВАК РФ и автоматическая классификация текстов по направлениям науки / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2018. № 144. С. 44-102. DOI 10.21515/1990-4665-144-033. EDN SNOCOD.
- 3. Луценко, Е. В. Синтез семантических ядер научных специальностей ВАК РФ и автоматическая классификации статей по научным специальностям с применением АСК-анализа и интеллектуальной системы "Эйдос" (на примере Научного журнала КубГАУ и его научных специальностей: механизации, агрономии и ветеринарии) / Е. В. Луценко, Н. В. Андрафанова, Н. В. Потапова // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2019. № 145. С. 31-102. DOI 10.21515/1990-4665-145-033. EDN YXHFVB.
- 4. Луценко, Е. В. Автоматизированный системно-когнитивный анализ в ветеринарии (на примере разработки диагностических тестов) / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2018. № 137. С. 143-196. DOI 10.21515/1990-4665-137-031. EDN YSTTRN.
- 5. Луценко, Е. В. Реализация тестов и супертестов для ветеринарной и медицинской диагностики в среде системы искусственного интеллекта "Эйдос-X++" без программирования / Е. В. Луценко, В. Е. Коржаков // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2013.- N 99.- C.99-140.- EDN TJAPAV.
- 6. Луценко, Е. В. Агломеративная когнитивная кластеризация нозологических образов в ветеринарии / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. − 2018. − № 138. − С. 122-139. − DOI 10.21515/1990-4665-138-033. − EDN XOQERN.
- 7. Луценко, Е. В. Агломеративная когнитивная кластеризация симптомов и синдромов в ветеринарии / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2018. № 139. С. 99-116. DOI 10.21515/1990-4665-139-033. EDN OURBBV.
- 8. Луценко, Е. В. Автоматизированный системно-когнитивный анализ и классификация пород крупного рогатого скота / Е. В. Луценко, Е. К. Печурина // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2018. № 142. С. 68-95. DOI 10.21515/1990-4665-142-033. EDN YOCDPN.

- 9. Луценко, Е. В. Автоматизированный системно-когнитивный анализ антибиотиков в ветеринарии / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. − 2018. − № 140. − С. 171-220. − DOI 10.21515/1990-4665-140-033. − EDN XTXJAT.
- 10. Луценко, Е. В. Автоматизированный системно-когнитивный анализ влияния пробиотиков в рационах на телосложение бычков / Е. В. Луценко, Е. К. Печурина // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2019. № 146. С. 68-93. DOI 10.21515/1990-4665-146-033. EDN PPTRDF.
- 11. Луценко, Е. В. Разработка ветеринарного теста для диагностики желудочнокишечных заболеваний лошади на основе данных репозитория UCI с применением АСК-анализа / Е. В. Луценко, Е. К. Печурина, А. Э. Сергеев // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. − 2018. − № 141. − С. 111-175. − DOI 10.21515/1990-4665-141-033. − EDN YKWDLF.
- 12. Луценко, Е. В. Математическое и численное моделирование взаимосвязи морфологического, биохимического и микроэлементного состава крови бычков герефордской породы и их размеров / Е. В. Луценко, В. Г. Лежнев, Н. И. Ковелин // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2018. № 143. С. 49-88. DOI 10.21515/1990-4665-143-033. EDN YPENJZ.
- 13. Луценко, Е. В. Когнитивная информационно-измерительная квалиметрическая система для определения содержания жира и белка в коровьем молоке по параметрам тензиограмм динамического поверхностного натяжения на границе раздела молоко/воздух / Е. В. Луценко, Е. К. Печурина, А. Э. Сергеев // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2019. № 151. С. 138-192. DOI 10.21515/1990-4665-151-015. EDN RGKOFL.
- 14. Работы проф.Е.В.Луценко & С° по когнитивной агрономии и когнитивной ампелографии: http://lc.kubagro.ru/aidos/Works on cognitive agronomy.htm
- 15. Луценко, Е. В. Теория информации и когнитивные технологии в моделировании сложных многопараметрических динамических технических систем / Е. В. Луценко, Г. В. Серга // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2016. № 121. С. 68-115. DOI 10.21515/1990-4665-121-002. EDN WWSKQL.
- 16. Lutsenko E.V. Cognitive veterinary medicine // February 2023, DOI: 10.13140/RG.2.2.28032.92163, License CC BY 4.0, https://www.researchgate.net/publication/368476220
 - 17. Сайт проф.Е.В.Луценко: http://lc.kubagro.ru/index.htm
 - 18. Блог проф.Е.В.Луценко: https://www.researchgate.net/profile/Eugene-Lutsenko
- 19. Луценко, Е. В. Автоматизированный системно-когнитивный анализ в управлении активными объектами : (системная теория информации и ее применение в исследовании экономических, социально-психологических, технологических и организационно-технических систем) / Е. В. Луценко. Краснодар : Кубанский государственный аграрный университет имени И.Т. Трубилина, 2002. 605 с. ISBN 5-94672-020-1. EDN OCZFHC.
- 20. Орлов, А. И. Системная нечеткая интервальная математика / А. И. Орлов, Е. В. Луценко. Краснодар : Кубанский государственный аграрный университет, 2014. 600 с. ISBN 978-5-94672-757-0. EDN RZJXZZ.

- 21. Луценко, Е. В. Метризация измерительных шкал различных типов и совместная сопоставимая количественная обработка разнородных факторов в системно-когнитивном анализе и системе "Эйдос" / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2013.- № 92.- C. 61-71.- EDN RNEGHR.
- 22. Работы проф.Е.В.Луценко по ACK-анализу текстов: http://lc.kubagro.ru/aidos/Works on ASK-analysis of texts.htm
- 23. Луценко, Е. В. Сценарный и спектральный автоматизированный системно-когнитивный анализ / Е. В. Луценко. Краснодар : Кубанский государственный аграрный университет им. И.Т. Трубилина, 2021. 288 с. DOI 10.13140/RG.2.2.22981.37608. EDN ZQLITW.
- 24. Франселла Ф., Баннистер Д. Новый метод исследования личности: Руководство по репертуарным личностным методикам: Пер. с англ. / Общ. ред. и предисл. Ю.М. Забродина и В.И. Похилько.— М.: Прогресс, 1987. 236 с.: ил. http://www.isras.ru/files/File/publ/Scan_2018/Novy_metod_issledovaniya_lichnosti_1987.pdf
- 25. Работы проф.Е.В.Луценко & С° по выявлению, представлению и использованию знаний, логике и методологии научного познания: http://lc.kubagro.ru/aidos/Works_on_identification_presentation_and_use_of_knowledge.htm
- 26. Луценко Е.В. Моделирование сложных многофакторных нелинейных объектов управления на основе фрагментированных зашумленных эмпирических данных большой размерности в системно-когнитивном анализе и интеллектуальной системе «Эйдос-Х++» / Е.В. Луценко, В.Е. Коржаков // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. Краснодар: КубГАУ, 2013. №07(091). С. 164 188. IDA [article ID]: 0911307012. Режим доступа: http://ej.kubagro.ru/2013/07/pdf/12.pdf, 1,562 у.п.л.

Научное издание

Луценко Евгений Вениаминович

АВТОМАТИЗИРОВАННЫЙ СИСТЕМНО-КОГНИТИВНЫЙ АНАЛИЗ В ВЕТЕРИНАРИИ

Монография

В авторской редакции Компьютерная верстка – Е. В. Луценко Макет обложки – Е. В. Луценко

Подписано в печать 25.02.2023. Формат $60 \times 84^{-1}/_{16}$. Усл. печ. л. -9,4. Уч.-изд. л. -7,4. Тираж 500 экз. Заказ № 50 экз.

Типография Кубанского государственного аграрного университета. 350044, г. Краснодар, ул. Калинина, 13