

УТВЕРЖДАЮ:

Врио директора

Государственного учреждения

«Институт проблем искусственного интеллекта» (ГУ ИПИИ),

**Вербина Полина Викторовна**



« 15 » октября 2019 г.

## ОТЗЫВ

ведущей организации на диссертацию Бурлаевой Екатерины Игоревны на тему: «Совершенствование методов системного анализа в задачах управления и систематизации специализированной информации», представленную на соискание ученой степени кандидата технических наук по научной специальности 05.13.01 – «Системный анализ, управление и обработка информации (по отраслям) (технические науки)»

### 1. Актуальность темы диссертационной работы.

Важнейшей характеристикой современного этапа развития общества является информатизация всех сторон деятельности человека. Во всех отраслях промышленности, экономики, инфраструктуры циркулируют и используются обширные массивы информации, в большинстве своём находящиеся в несистематизированной текстовой форме, что затрудняет получение требуемых сведений по конкретной тематике.

Это вызывает острую необходимость развития работ по созданию систем анализа, систематизации и управления специализированной текстовой информацией, так как даже высококвалифицированные эксперты испытывают затруднения по организации поиска документов и распределению полученных текстовых данных по тематикам.

Одним из способов систематизации, управления данными и их анализа является классификация информации, состоящая из сортировки текстовых документов по заранее определенным категориям. В этой связи тема работы является, безусловно, актуальной.

### 2. Основные научные результаты, их значимость для науки и практики

2.1 Научная новизна и практическая значимость полученных результатов состоит в следующем.

#### Научная новизна:

– впервые предложена общая модель автоматической систематизации и управления информацией, основанная на объединении достоинств синтагматических и парадигматических подходов, использование

ГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ  
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ  
«ДОНЕЦКИЙ НАЦИОНАЛЬНЫЙ  
ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»  
Вх. № 16/102.  
« 24 » 10 20 19 г.

которой позволяет повысить полноту и точность работы модели в среднем на 32.5% и 31.5% соответственно;

- получила дальнейшее развитие и модернизирована модель классификации текстовой информации на основе внесения изменений в структуру алгоритма её построения, что позволяет повысить полноту и точность работы модели еще на 5.5% и 8.5 % соответственно;

- экспериментально обоснована модель вычислительной композиции распределения веса слов в текстовом документе; при разном распределении веса термина, повышение качества работы предложенной общей модели систематизации и управления информацией варьируется в пределах 10% в зависимости от используемой композиции.

#### Практическая значимость:

- повышена эффективность предложенной общей модели систематизации и управления текстовой информацией на основе разработанных модернизированных моделей классификации информации за счет внесения изменений в структуру алгоритма её построения и вычислительных композиций распределения веса слов в документах, правил отбора неинформативных признаков и способов взвешивания термов;

- применение предложенных усовершенствований обеспечивает дополнительное повышение качества распределения информации на 28%.

- предложенная практическая реализация усовершенствованной модели систематизации и управления информацией позволяет формировать текстовые базы данных классифицированной информации в автоматическом режиме.

#### 2.2. Ценность научных работ соискателя учёной степени состоит в следующем.

- Предложенная комплексная методика построения модели автоматической классификации и статистического анализа является совершенствованием существующих подходов к обработке информации и в дальнейшем может быть расширена и дополнена функциями автоматического и автоматизированного тематического анализа потоков текстовой информации для расширения количества тематик, по которым распределяются текстовые документы, а также повышением качества модели автоматической обработки информации.

- Структура статистических баз данных, формируемых с помощью предложенной технологии, позволяет ставить и решать большой спектр статистических и математических расчетных задач, и задач, связанных с принятием решений, имеющих место в информационных системах.

- Развитие данной разработки может осуществляться путем дополнения ее новыми решениями в области морфологического, синтаксического и семантического анализа языков, для усовершенствования методов управления и систематизации специализированной текстовой информации.

Практическое значение полученных результатов подтверждается внедрением в практику организации информационных массивов и баз

данных с целью совершенствования компьютерной технологии прогноза в отделе сдвижения земной поверхности и охраны подрабатываемых объектов (СЗПО) Республиканского академического научно-исследовательского и проектно-конструкторского института горной геологии, геомеханики, геофизики и маркшейдерского дела (РАНИМИ) (справка о внедрении № 01/140 от 20.05.2019 г.).

### **3. Рекомендации по использованию результатов и выводов диссертации**

Результаты исследований имеют широкий спектр применения для различных предметных областей:

- предложенная практическая реализация усовершенствованной модели систематизации и управления информацией позволяет формировать текстовые базы данных классифицированной информации в автоматическом режиме;

- на основании результатов классификации имеется возможность формировать аналитические задачи и статистические базы данных по результатам обработки текстов, автоматизировать работу специалистов-аналитиков, осуществляющих тематический анализ текстовой информации, и ведение аналитических задач в различных предметных областях, что может послужить функциональным дополнением и развитием информационных систем различных организаций.

### **4. Общие замечания**

1. В главе 2 излишнее внимание уделено формальному определению ряда известных терминов (морфология, семантика и др.), тогда как новые термины не всегда трактуются достаточно подробно (стемминг, парсер и др.).

2. Обоснование необходимости модернизации подходов к решению задачи классификации сопровождается анализом недостатков используемых в настоящее время способов, при этом не всегда достаточно подробно проводится сравнительный анализ.

3. В главе 2 описан ряд алгоритмов машинного обучения, при этом не рассматривается их компьютерная реализация.

4. В главе 3 основное внимание в работе сосредоточено на обосновании алгоритмов, вместе с тем практически не уделяется внимания развитию структур и техническому обеспечению систем.

5. В главе 4 указывается, что в предложенную модернизированную модель системы управления и систематизации информации были включены модификации, улучшившие ее работу, однако при этом не показано, какие улучшения имеют место.

6. В работе не уделено внимания решению задач экономического анализа в аспекте внедрения предложенных разработок.

## Заключение

В целом представленная диссертация является завершенной научно-исследовательской работой, содержащей новые теоретические и практические результаты, обладает научной новизной и практической значимостью, соответствует научной специальности 05.13.01 – «Системный анализ, управление и обработка информации (по отраслям) (технические науки)», отвечает требованиям п. 2.2 «Положения о присуждении ученых степеней», предъявляемым к кандидатским диссертациям, а ее автор, Бурлаева Екатерина Игоревна, за полученное новое решение актуальной научно-технической задачи, заключающейся в повышении эффективности управления и систематизации специализированной текстовой информации больших объёмов с учетом особенностей русскоязычных текстовых документов, реализованных с помощью модернизация моделей, методик, и алгоритмов системного анализа, искусственного интеллекта и компьютерной лингвистики, заслуживает присуждения ученой степени кандидата технических наук по специальности 05.13.01 – «Системный анализ, управление и обработка информации (по отраслям) (технические науки)».

Отзыв обсужден и одобрен на заседании отдела распознавания зрительных образов Государственного учреждения «Институт проблем искусственного интеллекта» (ГУ ИПИИ)

«15» 10 2019 г., протокол № 2 .

Заведующий отделом распознавания зрительных образов Государственного учреждения «Институт проблем искусственного интеллекта» (ГУ ИПИИ) Бондарчук Виктория Валерьевна

(личная подпись)

Кандидатская диссертация защищена по специальности 05.13.07 – Автоматизация технологических процессов и производств в горной промышленности.

Адрес: 283048, ДНР, г. Донецк, ул. Артема, 118-б

Тел.: +38 (062) 311-34-24

E-mail: [iai\\_sb\\_ivanova@mail.ru](mailto:iai_sb_ivanova@mail.ru);

<http://guiaidn.ru>

Я, Виктория Валерьевна Бондарчук, согласна на автоматизированную обработку персональных данных, приведенных в этом документе

В.В.Бондарчук

(личная подпись)



Подпись Бондарчук В.В. заверяю  
В.В. Вербина