



УТВЕРЖДАЮ

Проректор по научной работе и  
инновационно-коммуникационным  
технологиям

к.т.н. Матвеев С.А.

2015 г.

### ОТЗЫВ

на автореферат диссертационной работы Смирнова Сергея Владимировича по теме «Технология и система автоматической корректировки результатов при распознавании архивных документов», представленной на соискание ученой степени кандидата технических наук по специальности 05.13.11 — Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей.

#### Актуальность темы диссертации

Оцифровка фондов библиотек, музеев, архивов представляет собой активно развивающуюся область применения информационных технологий и средств искусственного интеллекта. Данные технологии и средства направлены на создание программных инструментов, обеспечивающих генерацию электронного образа документа, последующую его обработку, распознавание и корректировку. При этом не существует универсального средства, дающего оптимальные результаты при обработке документов различных видов, тематических областей и категорий качества. В каждом случае производится настройка и адаптация существующих или разработка новых программных средств для повышения точности и сокращения количества допускаемых ошибок. Актуальной задачей является автоматизация процесса выбора оптимальной конфигурации и корректировки получаемых результатов.

#### Научная и практическая значимость полученных результатов

Ценность диссертационной работы Смирнова С.В. заключается в том, что разработанные технология и система распознавания разнотипных архивных материалов носят достаточно универсальный характер и могут быть применены в смежных отраслях. Они структурируют весь процесс обработки от загрузки электронного образа до получения текстового представления и предоставляют инструментарий ограничения пространства конфигурации для наиболее эффективной обработки того или иного вида документов. Также в работе существенное внимание уделено прикладному аспекту: разработке метода автоматической

корректировки результатов оптического распознавания, который учитывает контекст документа и производит автоматическую самоадаптацию под специфику используемого лексикона.

О возможности практического применения результатов работы можно судить по успешным результатам обработки документов центральных государственных архивов Санкт-Петербурга, представленным в главе 4. Очевидно, что применение полученных разработок позволяет существенно сократить объем дорогостоящей ручной работы и увеличить скорость получения качественного представления бумажных документов в электронной форме.

### **Замечания по содержанию автореферата**

1. В автореферате говорится о решении задачи отбора корректировок методом анаграмм, но не приводится его описания, что несколько препятствует детальному пониманию алгоритма автоматической корректировки результатов распознавания.
2. Отсутствуют примеры генерируемых структур данных на различных этапах корректировки, их наличие упростило бы понимание и внесло дополнительную наглядность.

### **Заключение**

Приведенные замечания не оказывают влияния на значимость представленных в работе результатов. В целом, судя по автореферату, диссертационная работа выполнена на высоком научно-техническом уровне и удовлетворяет требованиям ВАК РФ, предъявляемым к кандидатским диссертациям, а ее автор Смирнов Сергей Владимирович заслуживает присвоения ученой степени кандидата технических наук по специальности 05.13.11 — «Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей».

Отзыв составил:

Доцент кафедры И9 "Систем управления и компьютерных технологий" Федерального государственного бюджетного образовательного учреждения высшего профессионального образования «Балтийский государственный технический университет "ВОЕНМЕХ" им. Д.Ф. Устинова»

Доцент, к.т.н. Гуцин Артем Николаевич



«28» апреля 2015 г.