

DOI: 10.25728/avtprom.2024.05.04

А.С. Сержантов, Г.Д. Волкова (ФГБОУ ВО «МГТУ «СТАНКИН»)

Анализ подходов к извлечению знаний из конструкторской документации

Рассматривается проблема извлечения и представления знаний в технической документации. Исследованы особенности представления знаний в технической документации, оформленной в виде альбомов нормативно-справочных документов. Выделены критерии для анализа подходов к извлечению знаний. Сформированы требования к подходам, решающим задачу извлечения знаний. Проведен аналитический обзор существующих подходов к извлечению знаний. Дана оценка применимости существующих подходов к извлечению знаний из технической документации.

Ключевые слова: извлечение знаний, предметные знания, техническая документация, альбомы технических решений, обработка естественного языка, семантическая сеть, автоматизация предметных задач.

*Сержантов Артем Сергеевич – аспирант,
Волкова Галина Дмитриевна – д-р техн. наук, проф., ФГБОУ ВО "МГТУ "СТАНКИН".*

Список литературы

- 1. Волкова Г.Д. Концептуальное моделирование при создании прикладных автоматизированных систем // Тр. V национальной конференции с международным участием "Искусственный интеллект-96". КИИ-96 – Казань: Ассоциация искусственного интеллекта, 1996г., т. III, с. 440-446.*
- 2. Смоляков А.Л. Извлечение знаний из текстовой информации с помощью метода шаблонов // Вестник С.-Петербургского университета. Прикладная математика. Информатика. Процессы управления. 2008. № 2. С. 44-50.*
- 3. Ondrejcek M. et al. Information extraction from scanned engineering drawings // National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign, Image Spatial Data Analysis Group. 2009.*
- 4. Наместников А.М., Филиппов А.А., Субхангулов Р.А. Разработка инструментария для интеллектуального анализа технической документации // Известия Самарского научного центра РАН. 2011. №4-4.*
- 5. Орлова Ю.А. Методика анализа текста технического задания // Известия ТулГУ. Технические науки. 2011. №3.*
- 6. Souili A., Cavallucci D., Rousselot F. Natural language processing (NLP) – a solution for knowledge extraction from patent unstructured data // Procedia engineering. 2015. Т. 131. С. 635-643.*
- 7. Диковицкий В.В., Шишаев М. Г., Пимешков В.К. Метод автоматизированного извлечения понятий и парадигматических отношений тезауруса из текстов на естественном языке на базе лексико-синтаксических шаблонов // Труды Кольского научного центра РАН. 2019. №9 (9).*
- 8. Kruit B., Boncz P., Urbani J. Extracting novel facts from tables for knowledge graph completion // The Semantic Web – ISWC 2019: 18th International Semantic Web Conference, Auckland, New Zealand. 2019. Proceedings, Part I 18. Springer International Publishing, 2019. С. 364-381.*
- 9. Agarwal S., Atreja S., Agarwal V. Extracting procedural knowledge from technical documents // arXiv preprint arXiv:2010.10156. 2020.*
- 10. Муромцев Д.И., Шилин И.А., Плюхин Д.А., Баймуратов И.Р. и др. Построение графов знаний нормативной документации на основе семантического моделирования и автоматического извлечения терминов // Научно-технический вестник информационных технологий, механики и оптики. 2021. №2.*

Serjantov A.S., Volkova G.D. Comparing the approaches and methods of knowledge retrieval from technical documentation

The problem of knowledge retrieval from technical documentation is examined. The paper investigates the features of technical knowledge representation in technical documentation issued as books of technical solutions. Criteria for analyzing the approaches to and methods of data retrieval are identified. Existing data retrieval techniques are overviewed and analyzed, their applicability to data retrieval from technical documentation is assessed.

Keywords: *knowledge retrieval, subject knowledge, technical documentation, books of technical solutions, natural language processing, semantic network, automation of subject tasks.*