

АВТОМАТИЗИРОВАННАЯ СИСТЕМА ОПТИМИЗАЦИИ ПРОЦЕССА НАЙМА И БЕЗОПАСНОГО УПРАВЛЕНИЯ ДАННЫМИ ПЕРСОНАЛА

Войтович Е.А.

*Белорусский государственный университет информатики и радиоэлектроники,
г. Минск, Республика Беларусь*

Научный руководитель: Бобровнича М.А. – ст. преподаватель кафедры ИПиЭ

Аннотация. В статье рассматривается автоматизированная система оптимизации процесса найма, которая является частью HRM-платформы. Система предназначена для ускорения первичного отбора кандидатов, повышения качества сопоставления резюме с требованиями вакансий и обеспечения контролируемой обработки персональных данных. В основе решения лежит применение методов обработки естественного языка (NLP) и больших языковых моделей для извлечения структурированных сведений из резюме, нормализации навыков, опыта и других значимых характеристик кандидата, а также для формирования объяснимого ранжирования. С практической точки зрения, система поддерживает резюме в форматах PDF и DOCX, сохранение связи извлечённых фактов с исходными фрагментами текста и использовании этих фрагментов при объяснении результатов анализа. Отдельное внимание уделяется обработке русскоязычных и англоязычных резюме, а также вопросам оценки качества AI-модуля и снижения рисков при работе с чувствительными данными.

Ключевые слова: HRM, найм персонала, скрининг, персональные данные, безопасность, резюме, PDF, DOCX, NLP, LLM, двуязычная обработка, объяснимость.

Введение. На этапе первичного отбора HR-специалисты обрабатывают большое количество резюме, представленных в разных форматах. Это усложняет сопоставление кандидатов с требованиями вакансии и требует значительных временных затрат. Стоит отметить, что одинаковые по смыслу навыки могут описываться разными формулировками, сведения об опыте работы – находиться в разных разделах, а важная информация – включаться в описание опыта или обязанностей. В итоге даже при наличии точных требований первичный скрининг остаётся трудоёмким [1].

Сложности усиливаются при обработке резюме в различных форматах (PDF и DOCX), а также при работе с русскоязычными и англоязычными документами. Кроме того, кадровые процессы требуют соблюдения норм защиты персональных данных, поскольку сведения о кандидатах относятся к чувствительной информации и должны обрабатываться в контролируемом режиме. Таким образом, разработка автоматизированной системы, позволяющей ускорить ранние этапы подбора, унифицировать предварительный анализ резюме и обеспечить безопасное управление кадровыми данными, является актуальной. Цель работы – описание автоматизированной системы оптимизации процесса найма, основанной на методах NLP и языковых моделях, с поддержкой объяснимого сопоставления кандидатов и вакансий и контролируемой обработки персональной информации [2].

Основная часть. Резюме представляет собой неструктурированный текст, в котором значимые для отбора сведения могут быть выражены по-разному и зависеть от контекста. По этой причине в основе системы лежит следующий процесс обработки документов, включающий извлечение текста из файлов PDF и DOCX, его очистку и последующую нормализацию. На данном этапе устраняются последствия форматирования, приводятся к общему формату даты, названия технологий и должностей, а также выделяются основные содержательные блоки: опыт работы, навыки, проекты и образование.

Результатом обработки является очищенный текст документа, и унифицированный профиль кандидата, который может использоваться повторно при изменении требований вакансии или правил оценки. Для кадровых процессов особенно важна возможность проверки полученных выводов, поэтому при извлечении фактов сохраняется связь с исходными фрагментами резюме – предложениями или абзацами, в которых подтверждаются навыки, должности или проектный опыт. Эти фрагменты используются при создании объяснения

результата, делает рекомендации системы более понятными для HR-специалиста.

Разрабатываемая система не рассматривается как замена HR-специалиста. Её назначение состоит в том, чтобы упростить первичный отбор и сделать его более последовательным. Для практического использования важна не только итоговая оценка соответствия кандидата вакансии, но и возможность увидеть, какие требования подтверждаются текстом резюме, какие отражены частично, а какие не подтверждены в явном виде. Такой подход позволяет сократить время обработки откликов, уменьшить нагрузку на специалистов по подбору персонала и сделать предварительное ранжирование кандидатов более обоснованным.

Поддержка двуязычного режима требует, чтобы система корректно извлекала сведения как из русскоязычных, так и из англоязычных резюме, не смешивая терминологию и не искажая смысл при приведении данных к единому виду. Для этого определяется язык документа, после чего применяются соответствующие правила нормализации и словарные соответствия. Сопоставление содержания выполняется не на уровне буквального перевода, а с использованием семантических представлений, позволяющих учитывать, близкие по смыслу формулировки. При этом исходные фрагменты текста сохраняются как подтверждение извлечённых фактов, а в унифицированном профиле фиксируются нормализованные обозначения навыков, ролей и других значимых характеристик.

Это позволяет перейти от анализа отдельных резюме в свободной форме к сравнению унифицированных профилей кандидатов. Для каждого соискателя система может выделять сведения об образовании, профессиональном опыте, ключевых навыках, проектах и других характеристиках, имеющих значение для конкретной вакансии. За счёт этого HR-специалисту проще сопоставлять кандидатов по единым критериям, не тратя дополнительное время на интерпретацию однотипной информации, представленной в разной форме. Особенно это важно при большом количестве откликов, когда необходимо сохранить и скорость первичного отбора, и достаточную обоснованность предварительных решений.

Сопоставление кандидата с вакансией строится на сочетании формализованных требований и оценки семантической близости текстовых описаний. Вакансия задаёт критерии в виде набора компетенций, требований к опыту и дополнительных условий, тогда как резюме представляется как совокупность нормализованных навыков, ролей, предметных областей и подтверждающих фрагментов опыта. На выходе формируется ранжирование кандидатов и пояснение, показывающее, какие требования подтверждаются содержанием резюме, какими фрагментами это подтверждается и какие дефициты компетенций повлияли на итоговую оценку. Качество работы модуля оценивается по метрикам точности и полноты извлечения сущностей, а также по метрикам ранжирования, например NDCG и MRR. Дополнительно учитывается устойчивость системы к незначительным переформулировкам текста резюме и вакансии [3].

Также существует вопрос единообразия оценки кандидатов на начальном этапе подбора. Хотя автоматизированная система не устраняет полностью риск смещений, она позволяет сделать первичный отбор более последовательным за счёт использования заранее заданных критериев сопоставления вакансии и резюме. Важным преимуществом этого подхода является прозрачность предварительных решений: HR-специалист может видеть, какие характеристики кандидата были учтены как релевантные, а какие оказались выражены недостаточно явно. Это повышает доверие к результатам анализа и упрощает проверку спорных случаев.

Безопасное управление персональными данными является обязательной частью системы. Кадровая информация – чувствительные данные и требует контролируемой обработки. В данной платформе это достигается за счет разграничения прав доступа к резюме и результатам анализа, фиксации действий пользователей и ограничении объёма данных, доступных на каждом этапе работы. Такой подход позволяет снизить риски, связанные кадровыми процессами, и повысить безопасность обработки персональной информации [4].

Практическая ценность подобных решений определяется не только точностью сопоставления резюме и вакансии, но и возможностью их включения в реальный кадровый процесс. Система первичного отбора должна автоматизировать рутинные операции, сокращать время предварительной обработки откликов и поддерживать единый порядок анализа кандидатов. Существенное значение при этом имеет поддержка резюме в форматах PDF и DOCX, а также учёт двуязычного контекста при обработке русскоязычных и англоязычных

документов. Таким образом, автоматизированный анализ резюме может использоваться как практический инструмент повышения эффективности подбора персонала.

Заключение. В работе рассмотрена автоматизированная система оптимизации процесса найма, ориентированная на ускорение первичного отбора кандидатов и повышение обоснованности предварительного ранжирования резюме. Использование методов обработки естественного языка и языковых моделей позволяет извлекать из резюме структурированные сведения, приводить их к унифицированному представлению и сопоставлять с требованиями вакансии с учётом семантической близости формулировок. Это особенно важно при обработке документов в форматах PDF и DOCX, а также при работе с русскоязычными и англоязычными резюме.

Данная платформа несет следующую практическую значимость: HR-специалист получает не только итоговую оценку соответствия кандидата вакансии, но и основания для такой оценки в виде фрагментов исходного текста резюме. Такой подход позволяет сделать первичный скрининг более единообразным, сократить время обработки откликов и повысить прозрачность предварительных решений. Также система подразумевает учёт требований к защите персональных данных за счёт разграничения доступа к информации и контролируемой обработки кадровых документов, что повышает надёжность использования системы в реальных HR-процессах [5].

Список литературы

1. Особенности подбора персонала: интеллектуальный анализ текстов резюме и вакансий // CyberLeninka. – URL: cyberleninka.ru/article/n/osobennosti-podbora-personala-intellektualnyy-analiz-tekstov-rezume-i-vakansiy.
2. Закон Республики Беларусь от 07.05.2021 № 99-З «О защите персональных данных» // Национальный правовой Интернет-портал Республики Беларусь. – URL: pravo.by/document/?guid=12551&p0=H12100099.
3. Ранжирование резюме с использованием NLP (выпускная квалификационная работа) // НИУ ВШЭ. – URL: hse.ru/edu/vkr/1052911017.
4. Приказ Оперативно-аналитического центра при Президенте Республики Беларусь от 12.11.2021 № 195 «О технической и криптографической защите персональных данных» // Национальный правовой Интернет-портал Республики Беларусь. – URL: pravo.by/document/?guid=3961&p0=T62104893.
5. Национальный центр защиты персональных данных Республики Беларусь. Правовые основания обработки персональных данных (разъяснение). – URL: cpd.by/zachita-personalnyh-dannyh/operatoru/pravovye-osnovaniya-obrabotki/.

UDC 004.8:331.108.3

AUTOMATED SYSTEM FOR HIRING PROCESS OPTIMIZATION AND SECURE PERSONAL DATA MANAGEMENT

Vaitovich Y.A.

Belarusian State University of Informatics and Radioelectronics, Minsk, Republic of Belarus

Bobrovnichaya M.A. – Senior Lecturer of the Department of IPiE

Annotation. The paper presents an automated system for hiring process optimization developed as a part of an HRM platform. The system is aimed at accelerating the initial screening stage, improving the consistency of candidate-to-vacancy matching, and supporting secure processing of personal data. The proposed solution applies natural language processing methods and large language models to extract structured information from CVs, normalize skills and experience, and generate an explainable ranking of candidates. The system supports PDF and DOCX resumes, preserves links between extracted facts and source text fragments, and operates with both Russian- and English-language CVs. Special attention is paid to evaluation of extraction and ranking quality, robustness of the AI module, and controlled handling of sensitive personnel data.

Keywords: HRM, hiring, personal data, security, CV, PDF, DOCX, NLP, LLM, bilingual processing, explainability.