Обработка опросных листов КИПиА с использованием ИИ-агента

В статье представлен программный ИИ-агент для обработки входящих опросных листов на приборы КИПиА. Система извлекает параметры из PDF-, Excel-файлов, многостраничных документов, сопоставляет их с внутренним кодировщиком, формирует код прибора и подготавливает черновик технико-коммерческого предложения. Для каждого значения сохраняется источник, указывается уровень уверенности, а выпуск ТКП возможен только после проверки инженером. Приведены результаты обработки заявок. Рассмотрены ситуации, в которых системе требуется контроль со стороны специалиста.

Агенты 360, г. Москва

Опросный лист как исходные данные для подбора прибора

Подготовка коммерческого предложения на контрольно-измерительные приборы (КИП) начинается с обработки опросного листа (ОЛ). В нем можно найти информацию об условиях работы прибора: измеряемую среду, диапазон и требуемую точность, способ присоединения к процессу, тип фланца, требования к взрывозащите и выходному сигналу. Задача инженера – извлечь эти сведения, сопоставить с каталогом, определить исполнение прибора, собрать его код и наконец подготовить технико-коммерческое предложение (ТКП). На первый взгляд, последовательность выглядит вполне формализованной. Но на деле разбор исходного документа занимает много времени.

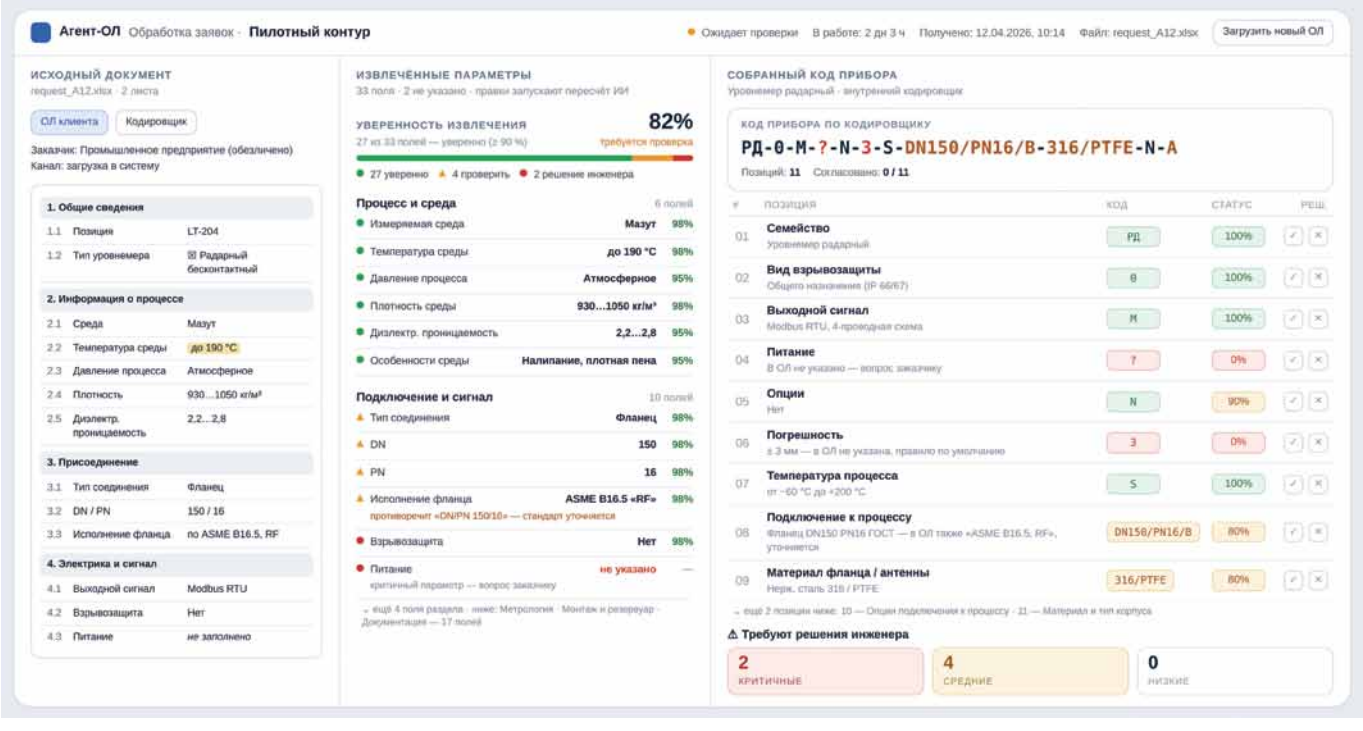


Рис. 1. Экран обработки заявки: исходный ОЛ, извлеченные параметры и собранный код прибора (данные условные)

Сложность заключается в том, что формат и содержание опросных листов меняются от заявки к заявке. В качестве ОЛ заказчик может прислать стандартный PDF-бланк, корпоративную форму Excel, многостраничный документ с общими требованиями, где нужные параметры надо найти в таблицах и дополнительных условиях. Встречаются и двуязычные документы. Например, в опросном листе могут одновременно стоять «DN/PN 150/16» и «фланец по ASME B16.5, RF» – обозначения из разных стандартов присоединения с разными размерами. Для инженера различие очевидно, но для системы автоматизации, которая извлекает данные из опросных листов, задача сложнее: распознать стандарт, заметить противоречие и не потерять его при формировании кода. Шаблоны, регулярные выражения и табличные парсеры хорошо использовать при стабильной структуре документа, однако если документы разные, каждое изменение требует доработки.

От входящего письма к проекту ТКП

Обработка опросных листов с помощью ИИ-агента разделяется на несколько последовательных операций. Сначала письмо и вложения приводятся к единому внутреннему представлению. При этом система сохраняет происхождение фрагмента: страницу, ячейку или область исходного документа. Затем извлекаются параметры. Вместе со значением сохраняются единицы измерения, ссылка на исходный фрагмент, а также уровень уверенности, который ИИ присваивает извлеченной информации. На следующем этапе параметры передаются во внутренний кодировщик, который формирует код исполнения прибора. Потом создается черновик ТКП, а спорные места отмечаются для проверки. Последний этап выполняет инженер: он подтверждает выбранные значения или вносит исправления, которые фиксируются в аудит-логе.

Интерфейс обработки заявки разделен на три области (рис. 1). Слева отображен исходный документ заказчика, в центре – извлеченные параметры с указанием уровня уверенности, справа – собранный код прибора. При выборе параметра соответствующий фрагмент исходного ОЛ подсвечивается. Поэтому инженер может проверить не только результат, но и место, из

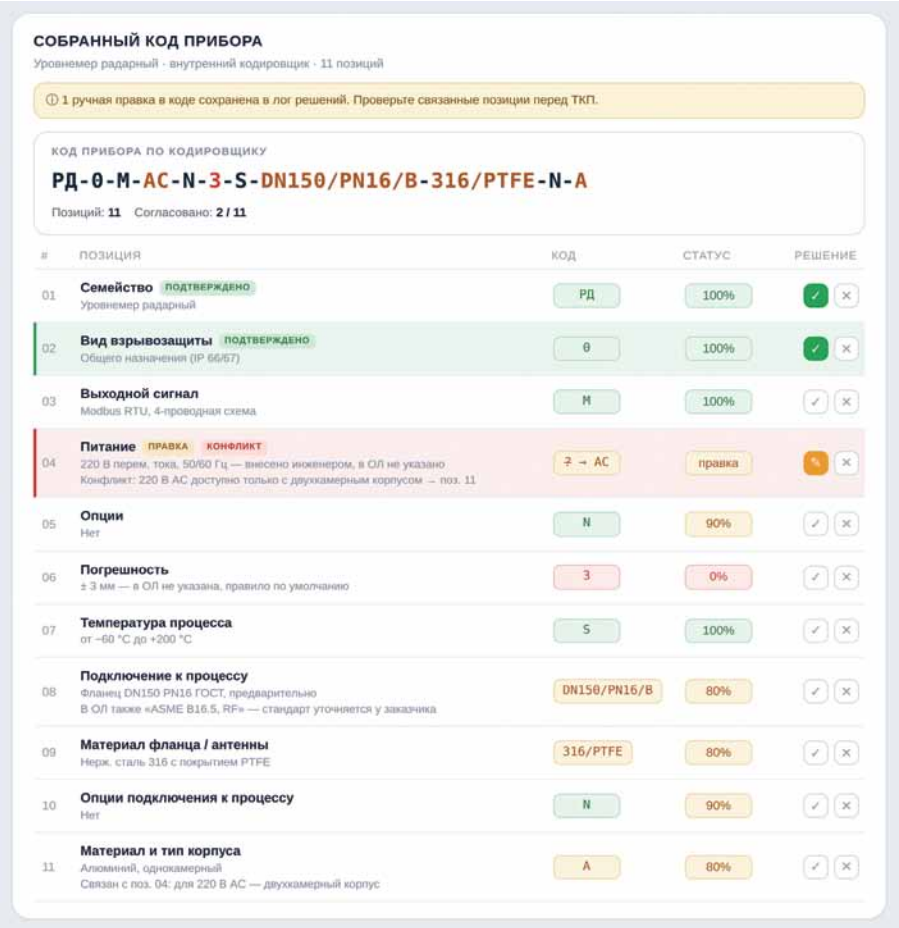


Рис. 2. Панель сборки кода прибора: правка инженера и связанный с ней конфликт помечены (данные условные)

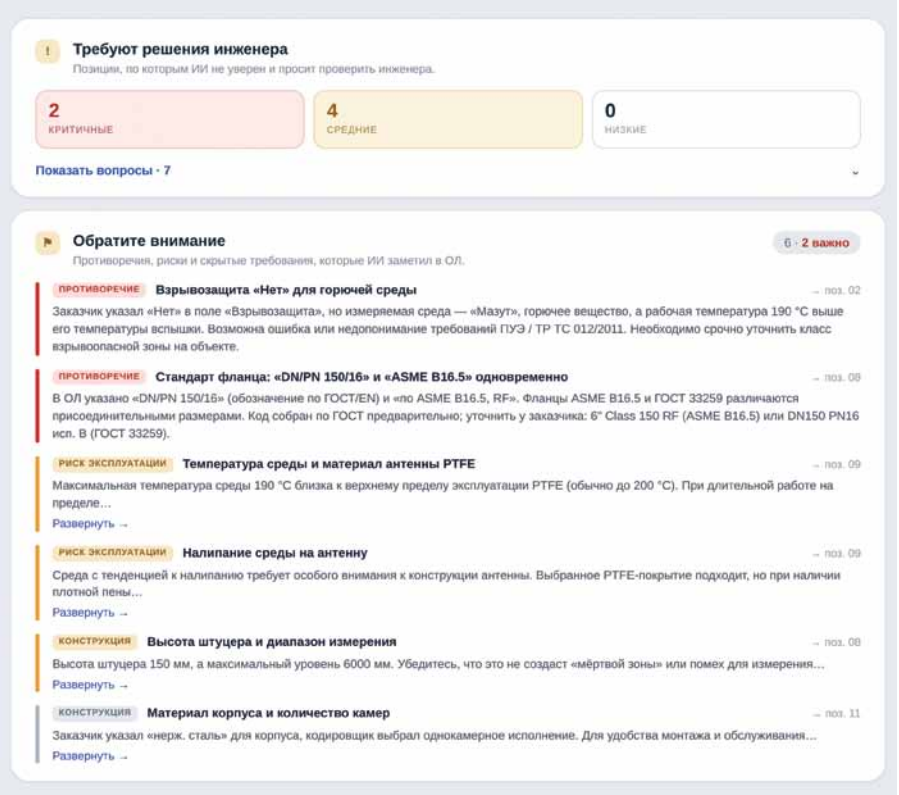


Рис. 3. Блок противоречий и скрытых требований, автоматически найденных в опросном листе (данные условные)

которого система получила исходное значение.

Формирование кода прибора и проверка параметров

После извлечения параметров система формирует код прибора по внутреннему кодировщику и показывает его по позициям (рис. 2). Для каждой позиции отображаются выбранный код и уровень уверенности. Если исходный документ не позволяет однозначно определить значение, система может предложить вариант в соответствии с логикой кодировщика или принятым правилом, но помечает такую позицию как требующую внимания. Инженер может подтвердить или отклонить выбор. При отклонении указывается краткое обоснование, которое сохраняется в журнале. В результате окончательное решение остается за специалистом, а история внесенных изменений не теряется.

Контроль ошибок и противоречий в исходных данных

В системе предусмотрено несколько уровней контроля. Для каждого параметра сохраняется ссылка на исходный фрагмент. Кроме того, параметры распределяются по степени риска. Критичными считаются, в частности, взрывозащита, температура процесса, измеряемая среда и питание. Пока инженер не подтвердит их значения, формирование ТКП блокируется. Для параметров с повышенным, но не критичным риском система только рекомендует дополнительную проверку.

Еще одна важная функция – поиск противоречий внутри опросного листа (рис. 3). Система может обнаружить, например, сочетание требования «взрывозащита – нет» с указанием горючей измеряемой среды. Другой пример – температура среды, близкая к предельно допустимой для выбранного материала. Такие ситуации выводятся в отдельный блок, чтобы инженер мог проверить их до формирования предложения.

Если исходных данных недостаточно, агент не должен самостоятельно восполнять критичный параметр. Но он формирует ряд конкретных вопросов заказчику. Это отличается от общего запроса: «Пришлите недостающие данные». Система указывает, какие именно сведения нужны для продолжения подбора. Все решения и исправления инженера сохраняются в аудит-логе: история позволяет восстановить, как формировалось конкретное ТКП, и использовать накопленные правки в процессе развития системы.

Результаты обработки заявок

Для проверки работы системы использовалась выборка реальных опросных листов разной сложности: стандартные PDF-бланки, фирменные Excel-формы, многостраничные документы требований и двуязычная заявка на фланцы с дюймовыми размерами по стандарту ASME. За эталон были взяты ТКП, подготовленные инженерами вручную по тем же исходным документам.

По данным авторов статьи, доля выбранных ИИ-агентом кодов приборов, совпавших с выбором человека, составила около 86 %. Один опросный лист система обрабатывала за 2–3 минуты, тогда как ручная подготовка занимала десятки минут. Проверяя результаты, разработчики внимательно искали расхождения, которые система вообще не обнаружила и не отметила снижением уверенности или флагом дополнительной проверки (поскольку система может ошибиться, но все ее ошибки должны быть видны). На использованной выборке таких «тихих» ошибок не обнаружено.

Где автоматизация пока ограничена

Наиболее сложной информацией для обработки остаются нетекстовые фрагменты ОЛ: чертежи, эскизы, сканы и фотографии бумажных бланков. Они распознаются хуже документов со структурированным текстом. Система не пытается автоматически довести такие участки до результата любой ценой: она снижает уровень уверенности и передает фрагмент инженеру. В итоге часть работы сохраняется за специалистом, но проблема становится явной и не превращается в незамеченную ошибку.

Эффективность системы заметна уже на этапе первичной обработки заявки. ИИ-агент позволяет быстро определить, есть ли в линейке подходящее исполнение прибора, и не тратить время на запросы, которые заведомо не соответствуют возможностям оборудования. Одновременно сокращается время подготовки первого ответа заказчику. На этом же этапе осуществляется автоматическая проверка противоречий в исходном документе: при ручной обработке такие несоответствия могут остаться незамеченными.

При этом искусственный интеллект не исключает из процесса инженера, его задача – сократить затраты труда и времени. Меняется распределение работы: система берет на себя первичный разбор документов, сопоставление параметров и подготовку варианта, а специалист проверяет критичные значения и принимает окончательное решение. Для задач, где цена ошибки высока, схема, сохраняющая источник каждого параметра и требующая подтверждения, выглядит более практичной, чем полностью автоматическая генерация ТКП.

Литература

1. *Vaswani A., Shazeer N., Parmar N. et al.* Attention Is All You Need // Advances in Neural Information Processing Systems. 2017. Vol. 30.
2. *Lewis P., Perez E., Piktus A. et al.* Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks // Advances in Neural Information Processing Systems. 2020. Vol. 33.
3. *Wu X., Xiao L., Sun Y. et al.* A Survey of Human-in-the-loop for Machine Learning // Future Generation Computer Systems. 2022. Vol. 135.

Ю. Г. Шаляпина,
В. Ф. С. Эспиноса,
Агенты 360, г. Москва,
тел.: +7 (909) 995-30-89,
эл. почта: j.shaliapina@mail.agenty360.ru,
сайт: agenty360.ru

Иллюстрации предоставлены ИП Шаляпиной Ю. Г.