

**Техническая спецификация
на закуп услуг по оцифровке архивных документов на бумажных носителях и
дальнейший экспорт в информационный ресурс Электронный архив**

1. Общие сведения

Выполнение работ по оцифровке бумажных документов ведомственного архива АО «НАК «Казатомпром».

1.1. Перечень документов, на основании которых проводится услуга

- Закон Республики Казахстан от 24 ноября 2015 года «Об информатизации»;
- Указ Президента Республики Казахстан от 9 декабря 1997 года №3787 «О формировании единого информационного пространства»;
- Закон Республики Казахстан от 22 декабря 1998 года № 326-І «О Национальном архивном фонде и архивах» (с изменениями и дополнениями по состоянию на 29.12.2014 г.);
- Приказ Министра культуры и спорта Республики Казахстан от 26 января 2015 года № 22 «Об утверждении Перечня типовых документов, образующихся в деятельности государственных и негосударственных организаций, с указанием сроков хранения».
- Другие нормативно-правовые акты РК, касающиеся регулирования вопросов информатизации и электронного документооборота в РК.

2. Характеристика объекта автоматизации

2.1. Краткие сведения об объекте автоматизации

АО "НАК "Казатомпром" является национальным оператором Республики Казахстан по импорту – экспорту урана, редких металлов, ядерного топлива для атомных электрических станций, специального оборудования, технологий и материалов двойного назначения.

Архив компании представляет собой структурированное хранилище бумажных документов, состоящее из 21000 дел (около 4 000 000 страниц).

2.2. Перечень объектов автоматизации

В рамках проекта необходимо произвести:

Оцифровку и перевод в электронный вид архивных документов Заказчика в единицах хранения: 15 515 дел (не более 3 103 000 страниц);

3. Оцифровка документов

Характеристика документов, подлежащих оцифровке.

№ п/п	Наименование	Значение	Удельный вес
1	Размеры листов	До А4	75%
		До А3	15%
		До А0	10%

2	Расположение информации	Одностороннее Двухстороннее	80% 20%
3	Возможность расшitia архивных дел	Возможно	100 %
4	Плотность бумаги	80 г/м и более Менее 80 г/м	95% 5%
5	Наличие ветхих документов	Имеется	5-6%
6	Тип информации	На русском языке На казахском языке	80% 20%

4. Условия оказания услуг

1. Оказание услуг должно осуществляться на территории Заказчика в отдельно выделенном помещении.
2. Заказчиком будет предоставлен представитель каждого структурного подразделения для передачи дел на сканирование, ведение журналов по выдаче и приемке документов, а также для консультаций по Номенклатуре дел каждого подразделения.
3. Поставщик должен владеть нормативными правовыми актами по документированию, управлению документацией и архивному делу.

5. Требования по сканированию

5.1. Требования к формату выходных данных

Метаданные должны быть переданы в формате XML, электронные копии документов в формате PDF/A.

5.2. Требование к создаваемым электронным копиям документов в комплексе

Сканирование документов должно проводиться с разрешением не менее 300 dpi в режиме 24 - bit в многостраничные файлы.

5.3. Требования к регламентам оцифровки

Оцифровку необходимо провести по следующей технологии:

1. Расшивка, разборка архивного дела;
2. Разделение дела по документам (для расшиваемых документов);
3. Проведение поточного сканирования для расшиваемых документов, проведение планшетного сканирования для не расшиваемых документов;
4. Проведение полнотекстового распознавания 100%;
5. Проведение верификации документов с более 10% не уверенно распознанных символов;
6. Проведение индексирования документов (не менее 10 атрибутивных полей к каждому типу документов);
7. Экспорт метаданных из Электронной копии документа (далее – ЭКД) в Электронный архив АО «НАК «Казатомпром»;
8. Сборка и сшивка документов в архивные дела.

После окончания работ по оцифровке должен быть создан массив метаданных и электронных копий документов, которые возможно будет найти и просмотреть исходя из прав доступа посредством Электронного архива АО «НАК «Казатомпром».

При проведении работ по атрибутированию должна использоваться следующая индексная информация:

При атрибутировании документов (xml)	При регистрации запросов в Системе ЭА
<p>- Хранилища.</p> <p>Категория ценностей документов (1 категория, 2 категория, 3 категория).</p> <p>Категория по срокам хранения (постоянного хранения, по личному составу, личные дела уволенных работников, временного (свыше 10 лет) хранения, документы на уничтожение, научно-техническая документация)</p>	<p>Характер запросов (тематический, социально-правовой).</p>
<p>- Сотрудники.</p> <p>Виды документов (управленческая документация, документы личного происхождения, научно-техническая документация, кинодокументы, видеодокументы, фотодокументы, фоно-документы).</p> <p>Типы документов (бумажные документы, видео-документы, фотодокументы).</p> <p>Категория документов (указы, приказы, постановления, распоряжение, протокола, решение, переписка, отчеты, паспорта сделок, договоры).</p>	<p>Категория запросов (административно-территориальное деление; создание, переименование, преобразование организаций, предприятий; акты ввода в эксплуатацию законченных строительством объектов; выделение земельных участков; трудовой стаж; зарплата; награждение; имущественные права (по документам нотариальных контор; другие).</p>
<p>Категория фонда (1 категория - областной, 2 категория - городской, 3 категория - районный, поселковый).</p>	<p>Формы собственности (частное, государ., негосудр.).</p>
<p>Основание поступления (истечение вед.срока хран, ликвидация учреждения с определением правопреемника, дарение, покупка, репарация, реституция.).</p>	<p>Организационно-правовая форма организации.</p>
<p>Изобразительные материалы (чертежи, карты, рисунки, фото, иное).</p>	<p>Виды именного каталога (Герои социалистического труда, лауреаты премий, заслуженные работники, почетные граждане, ветераны войны).</p>
<p>Собственность (федер., субъекта, улиц., физ.лиц, юр.лиц).</p>	<p>Страна запроса.</p> <p>Языки (казахский, русский, английский).</p>
<p>Вид архивного фонда (учреждения, организации; объединенный архивный фонд; л/п-личного происхождения; архивная коллекция).</p>	<p>Цель использования запроса (полит. народхоз., научное, соц-культ.).</p>
<p>Характеристика секретности (откр.- открытый, с- секретный, сч- частично секретный. дсп -для служебного пользования).</p>	<p>Характер ответа (Положит., отрицат.).</p>
	<p>Отрасли (государственная власть, конфессии)</p>

	объединения).
Доступ к документам (открытый, ограниченный).	Филиалы.
Способ воспроизведения (рукопись; машинопись; типографская; компьютерная распечатка; иное).	Виды обращений граждан (государственного управления, общественных и религиозных объединений, политических партий, коррупции государственного имущества, недвижимости, земли и землепользования, труда и занятости населения, брака и семьи, социального и пенсионного обеспечения). Срок исполнения запроса (до 15 дней, до 30 дн
Признаки группировки дел в описи (значимость документов; структурный; хронологический; функциональный).	
Подсистемы НТД (научная; проектная; конструкторская; технологическая; патентная).	
НСА (титульный лист, оглавление, предисловие, список сокращений, историческая справка).	

При работе со справочниками	
Справочник «Хранилища» содержит поля:	
<ul style="list-style-type: none"> - Номер хранилища; - Наименование хранилища; - Местоположение хранилища. 	
Справочник «Категория ценностей документов» содержит поля:	
<ul style="list-style-type: none"> - Наименование категории ценностей документов. 	
Справочник «Категория по срокам хранения» содержит поля:	
<ul style="list-style-type: none"> - Наименование категории по срокам хранения. 	
Справочник «Сотрудники» содержит поля:	
<ul style="list-style-type: none"> - ФИО сотрудника; - Должность сотрудника; - Состояние сотрудника; - Подразделение; - Признак руководителя подразделения. 	
Справочник «Виды документов» содержит поля:	
<ul style="list-style-type: none"> - Наименование вида документа. 	

- Краткое наименование вида документа.
Справочник «Типы документов» содержит поля: - Наименование типа документа. - Краткое наименование типа документа.
Справочник «Категория документов» содержит поля: - Наименование категории документов. - Краткое наименование категории документов.
Справочник «Характер запросов» содержит поля: - Наименование характера запроса.
Справочник «Категория запросов» содержит поля: - Наименование категории запросов. - Краткое наименование категории запросов.
Справочник «Формы собственности» содержит поля: - Наименование формы собственности. - Краткое наименование формы собственности.
Справочник «Организационно-правовая форма организации» содержит поля: - Наименование организационно-правовой формы организации. - Краткое наименование организационно-правовой формы организации.
Справочник «Виды именного каталога» содержит поля: - Наименование вида именного каталога. - Краткое наименование вида именного каталога.
Справочник «Страна запроса» содержит поля: - Наименование страны. - Наименование региона (СНГ, д/заруб.).
Справочник «Языки» содержит поля: - Наименование языка.
Справочник «Категория фонда» содержит поля: - Наименование категории фонда. - Описание категории фонда. Справочник «Основание поступления» содержит поля: - Наименование основания поступления. - Краткое наименование основания поступления.

<p>Справочник «Изобразительные материалы» содержит поля:</p> <ul style="list-style-type: none"> - Наименование изобразительного материала. - Краткое наименование изобразительного материала.
<p>Справочник «Собственность» содержит поля:</p> <ul style="list-style-type: none"> - Наименование собственности. - Краткое наименование собственности.
<p>Справочник «Вид архивного фонда» содержит поля:</p> <ul style="list-style-type: none"> - Наименование вида архивного фонда. - Краткое наименование вида архивного фонда.
<p>Справочник «Характеристика ограниченного доступа» содержит поля:</p> <ul style="list-style-type: none"> - Наименование характеристики ограниченного доступа. - Краткое наименование характеристики ограниченного доступа.
<p>Справочник «Доступ к документам» содержит поля:</p> <ul style="list-style-type: none"> - Наименование доступа к документам. - Краткое наименование доступа к документам.
<p>Справочник «Способ воспроизведения» содержит поля:</p> <ul style="list-style-type: none"> - Наименование способа воспроизведения. - Краткое наименование способа воспроизведения.
<p>Справочник «Признаки группировки дел в описи» содержит поля:</p> <ul style="list-style-type: none"> - Наименование признака группировки дел в описи. - Краткое наименование признака группировки дел в описи.
<p>Справочник «Подсистемы НТД» содержит поля:</p> <ul style="list-style-type: none"> - Наименование подсистемы НТД. - Краткое наименование подсистемы НТД.
<p>Справочник «НСА» содержит поля:</p> <ul style="list-style-type: none"> - Наименование НСА. - Краткое наименование НСА.
<p>Справочник «Цель использования запроса» содержит поля:</p> <ul style="list-style-type: none"> - Наименование цели использования запроса. - Краткое наименование цели использования запроса.
<p>Справочник «Характер ответа» содержит поля:</p> <ul style="list-style-type: none"> - Наименование характера ответа.

- Краткое наименование характера ответа.
Справочник «Отрасли» содержит поля:
- Номер отрасли.
- Наименование отрасли.
- Краткое наименование отрасли.
Справочник «ДЗО» содержит поля:
- Номер ДЗО.
- Наименование ДЗО.
- Местонахождение ДЗО.
Справочник «Виды обращений юридических и физических лиц» содержит поля:
- Наименование вида обращений юридических и физических лиц.
Справочник «Срок исполнения запроса» содержит поля:
- Наименование срока исполнения запроса.

5.4. Требования к функционированию программно-аппаратного комплекса

Процесс оцифровки должен представлять собой программно-аппаратный комплекс, который может функционировать, как в целом, так и частично.

Информационный обмен между компонентами комплекса должен осуществляться в рамках выделенной ЛВС. Протокол обмена- TCP/IP.

Аппаратно-программный комплекс ввода должен обеспечить ввод не менее 40 000 страниц документов в день.

5.5. Требования к технической надежности программно-аппаратного комплекса

Базовое ПО АРМ должно включать антивирусную программу, с актуальными базами сигнатур

5.6. Требования к вводу документов

Ввод документов в БД должен осуществляться с использованием ЛВС от АРМ виртуальной станции экспорта на сервер БД;

Комплекс ввода документов из архивов должен обеспечивать ввод документов, хранящихся как в расшитом, так и в не расшитом виде.

В состав комплекса ввода должно входить не менее 4 типов АРМ

АРМ сканирования расшитых документов;

АРМ сканирования не расшитых документов;

АРМ сканирования документов А0;

АРМ верификации и индексирования.

№	Наименование АРМ	Кол-во	
1	АРМ сканирования расшитых документов	Не менее 6	АРМ предоставляет Исполнитель на время выполнения работ по проекту, без передачи в собственность Заказчика. После окончания проекта оборудование возвращается Исполнителю.
2	АРМ сканирования не расшитых документов	Не менее 4	
3	АРМ сканирования документов А0	Не менее 2	
4	АРМ верификации и индексирования	Не менее 8	
	Общее количество АРМ	Не менее	

5.7. Требования к АРМ по сканированию расшитых документов

АРМ сканирования расшитых документов должен обеспечить возможность проведения следующих операций:

- поточного сканирования разделяемых на отдельные листы документов формата до А3;
- очистки, обрезку и улучшения качества изображения документов.

5.8. Требования к АРМ по сканированию не расшитых документов

АРМ сканирования не расшитых документов должен обеспечить возможность проведения следующих операций:

- сканирования документов формата до А3, которые невозможно разброшюровать, (расшить);
- очистки, обрезку и улучшению качества изображения документов.

5.9. Требования к АРМ по сканированию документов ф. А0

АРМ сканирования документов такого типа должен обеспечить возможность проведения следующих операций:

- сканирования документов формата до А0;
- очистки, обрезку и улучшению качества изображения документов.

5.10. Требования к АРМ по распознаванию, индексированию и верификации

АРМ должно обеспечить возможность присвоения электронной копии документа уникального набора значений (индексов), которые позволят однозначно идентифицировать документ в Электронный архив. Результатом индексирования должны стать наборы электронных образов и индексных полей, созданных в прикладных приложениях, соответствующие отдельной электронной копии и связанные со строкой в базе данных документов.

Для увеличения производительности залпового ввода в случае необходимости должна быть предусмотрена возможность масштабирования по числу АРМ.

5.11. Требования к показателям процесса проведения работ по оцифровке

Комплекс по созданию ЭКД должен содержать информацию о числе оцифрованных ЭКД по следующим показателям:

- за указанный период времени;
- по каждому из пользователей комплекса, проводящего работы по наполнению БД.

6. Требования к техническому обеспечению

Потенциальный Поставщик должен обладать парком оборудования, позволяющим осуществлять сканирование документов от А4 до 2А0 и представленным к работе в течении 5 дней после подписания Договора.

На всю указанную технику Поставщик должен предоставить копии бухгалтерских документов.

6.1. Аппаратные требования к сканирующей технике

6.1.1. Сканер для расшитых документов

Для сканирования расшитых документов необходимо использовать парк сканеров, следующих типов:

➤ Сканер расшитых документов – тип 1

Сканер должен обладать следующими характеристиками и соответствовать следующим минимальным требованиям:

Иметь протяжный тип сканирования;
Иметь надёжную, бережную работу с ветхими документами и документами плотностью менее 80 гр./см³;
Иметь формат вывода - черно-белый, оттенки серого, цветной;
Иметь высокую скорость сканирования – 100 стр./мин при 300 dpi при любом цветовом режиме;
Иметь ежедневную рекомендуемую нагрузку – 100 000 листов в день;
Иметь интеллектуальную систему предотвращения заторов бумаги, использующую контроль толщины каждого листа и многоуровневую систему определения двойного захвата - ультразвуковой датчик, контроль длины листа;
Иметь емкость автоподатчика – 500 листов;
Иметь дуплексное сканирование;
Сканировать документы размером не менее - 297X420 мм;
Иметь интерфейс - SCSI/USB;
Поддерживать стандарты - ISIS, TWAIN

Количество – не менее 2 единицы

➤ **Сканер расшитых документов – тип 2**

Сканер должен обладать следующими характеристиками и соответствовать следующим минимальным требованиям:

Иметь планшетно - протяжный тип сканирования;
Иметь надёжную, бережную работу с ветхими документами и документами плотностью менее 80 гр./см³;
Иметь формат вывода - черно-белый, оттенки серого, цветной;
Иметь высокую скорость сканирования – 90 стр./мин при 300 dpi при любом цветовом режиме;
Иметь ежедневную рекомендуемую нагрузку – 8 000 листов в день;
Иметь емкость автоподатчика – не менее 200 листов;
Иметь дуплексное сканирование;
Сканировать документы размером не менее – 297X420 мм;
Иметь интерфейс - USB 2.0/USCSI/UWSCSI;
Поддерживать стандарты - ISIS, TWAIN

Количество – не менее 2 единицы

➤ **Сканер расшитых документов – тип 3**

Сканер должен обладать следующими характеристиками и соответствовать следующим минимальным требованиям:

Иметь протяжный тип сканирования;
Иметь формат вывода - черно-белый, оттенки серого, цветной;
Иметь высокую скорость сканирования – 90 стр./мин при 300 dpi при любом цветовом режиме;
Иметь емкость автоподатчика – 500 листов;
Иметь дуплексное сканирование;
Сканировать документы размером не менее - 297X420 мм;
Иметь интерфейс - USB 2.0, SCSI;
Поддерживать стандарты - ISIS, TWAIN;
Иметь коррекцию перекоса, автоматическое определение формата страниц, пропуск пустых страниц, определение ориентации текста, предустановленные настройки гамма-коррекции, удаление цвета (красный, зелёный, синий / пользовательский цвет) / улучшение цвета (красный, зелёный, синий / пользовательский цвет), удаление точек, удаление полей, удаление следов перфорации, пакетное отделение, улучшение обработки контуров, предварительное сканирование, настройка области сканирования, технологию MultiStream, управляющие коды, отдельные настройки для лицевой и обратной сторон, ультразвуковое обнаружение двойной подачи.

Количество – не менее 2 единиц.

6.1.2 Сканер для не расшитых документов

Для сканирования не расшитых документов необходимо иметь следующий парк сканеров:

➤ Сканер не расшитых документов – тип 1

Сканер должен обладать следующими характеристиками и соответствовать следующим минимальным требованиям:

- Иметь планшетный тип сканирования;
- Иметь формат вывода - цветной: вход - 48-бит, выход - 24-бит, оттенки серого: вход - 16-бит, выход - 8-бит, ч/б: 1-бит;
- Иметь скорость сканирования – 2.48 сек (300dpi, A3, цветное сканирование) 2.10 сек (300 dpi, A3, оттенки серого и ч/б);
- Иметь ежедневную рекомендуемую нагрузку – 5 000 листов в день;
- Иметь максимальную площадь сканирования - 304,8X431,8 мм;
- Иметь интерфейс - USB 2.0;
- Поддерживать стандарты – TWAIN.
- Количество - не менее 3 единицы

➤ Сканер не расшитых документов – тип 2

Сканер должен обладать следующими характеристиками и соответствовать следующим минимальным требованиям:

- Иметь планшетный тип сканирования документов формата до A2;
- Иметь встроенную книжную колыбель с опциональным прижимным стеклом;
- Иметь формат глубины цвета – цветной - 24-бит;
- Иметь источник освещения - экономичную светодиодную подсветку (LED);
- Иметь время сканирования – 9 сек (300dpi, A2, цветное сканирование);
- Иметь ежедневную рекомендуемую нагрузку – 5 000 листов в день;
- Иметь размер области сканирования – 635X460 мм;
- Иметь интерфейс - Firewire 1394;
- Количество – не менее 1 единицы.

6.1.3. Сканер для документов формата A0

➤ Сканер для документов A0– тип 1

Сканер должен обладать следующими характеристиками и соответствовать следующим минимальным требованиям:

- Обеспечивать сканирование документов формата A0+;
- Обеспечивать скорость сканирования 305 мм/сек;
- Иметь ширину тракта 1420 мм;
- Иметь ширину области сканирования 1372 мм;
- Иметь сканирование без ограничения длины оригинала;
- Обеспечивать высокоскоростное сканирование с разрешением 1200 dpi;
- Иметь автоматическую систему настройки под толщину оригинала (AutomaticThicknessAdjustmentControl - ATAC) позволяющую сканировать оригиналы толщиной до 15 мм;
- Иметь специализированный привод протяжки оригинала (All-Wheel-Drive) обеспечивающий щадящий режим протяжки документа при сканировании и предохраняющий от повреждения ценные оригиналы.

Количество – не менее 1 единицы.

➤ Сканер для документов A0 – тип 2

Сканер должен обладать следующими характеристиками и соответствовать следующим минимальным требованиям:

- обеспечивать сканирование документов формата A2+
- обеспечивать скорость сканирования 7,62 см/сек (при монохромном сканировании –

25,4 см/сек);

- иметь ширину тракта 965 мм;
- иметь ширину области сканирования 914 мм;
- иметь сканирование без ограничения длины оригинала;
- обеспечивать высокоскоростное сканирование с максимальным разрешением 9600 dpi;
- иметь автоматическую систему настройки под толщину оригинала (AutomaticThicknessAdjustmentControl — АТАС) позволяющую сканировать оригиналы толщиной до 2 мм;
- иметь специализированный привод протяжки оригинала (All-Wheel-Drive) обеспечивающий щадящий режим протяжки документа при сканировании и предохраняющий от повреждения ценные оригиналы.

Количество – не менее 2 единицы.

6.2. Требования к серверной технике

Для внедрения Системы аппаратная платформа должна отвечать следующим требованиям:

На серверные оборудования должны быть установлены:

- ОС –Microsoft Windows Server Standard 2012;
- СУБД – Microsoft SQL Server Enterprise 2012.

Дополнительно должно быть выполнено требование по установке ПО:

1) на сервере СУБД:

- Операционная система Microsoft Windows Server Standard 2012;
- СУБД Microsoft SQL Server 2012 Enterprise;

2) на сервере приложений, интеграции и т.д.:

- операционная система Microsoft Windows Server Standard 2012;
- Web-сервер IIS 8.0;

Установка программного обеспечения Системы выполняется силами Поставщика.

Все оборудование и лицензии на данное серверное оборудование предоставляются Заказчиком.

6.3. Аппаратные требования АРМ

АРМ специалистов поставщика должны будут оснащены следующим компьютерным оборудованием (18 единиц) с характеристиками не ниже:

Системный блок:

- Оперативная память – 4 GB;
- Процессор - не ниже Core i5,
- 1 (один) жесткий диск – 500 GB, SATA, кэш 8MB, 7 200 RPM;
- Предустановленная операционная система Microsoft Windows.

Монитор:

- Диагональ - не менее 24"
- Разрешение-1920x1200 (16:10)

6.4. Аппаратные требования к графическим станциям

АРМ сканирования широкоформатных документов должны быть оснащены графическими станциями (3 единицы) с характеристиками не ниже:

Системный блок:

- Оперативная память – не менее 8 Гб;
- Процессор – не ниже Core i5, с тактовой частотой не менее 3,1 Гц;
- Жесткий диск – не менее 1 Тб;
- Видеокарта – GeForce или аналог, не менее 2 Гб;
- Предустановленная операционная система Microsoft Windows.

Монитор:

- Диагональ - не менее 24"
- Разрешение-1920x1200 (16:10)

7. Требования к сетям передачи данных

Ключевым требованием для функционирования системы является использование оптоволоконных каналов связи передачи данных. Требования к коммуникационному обеспечению:

- минимальная пропускная способность СПД между серверами и обслуживающим их персоналом – не менее 10 Мбит/с;
- постоянное VPN соединение;
- режим функционирования СПД – круглосуточный;
- реализуемые СПД должны иметь поддержку протоколов стека TCP/IP;
- наличие ЛВС на предприятиях, эксплуатирующих систему, с пропускной способностью не менее 10 Мбит/с и с доступом к магистральным каналам связи со скоростью передачи данных не менее 512 кбит/с;
- наличие связи между серверами, обеспечивающими работоспособность эксплуатируемой системы, по ЛВС со скоростью не менее 1 Гбит/с.

Минимальная пропускная способность канала передачи данных от клиента к серверу приложений для ЭА должна составлять 512 Кб.

Поставщик должен обеспечить нормальное функционирование всех пользователей Системы на необходимых каналах пропускной способностью 512 Кб и более.

Система ЭА должна поддерживать программно-аппаратный комплекс или ПО по оптимизации и балансировки сети передачи данных.

8. Требования к рабочим станциям

В целях корректной работы пользователей в ЭА, рабочие станции Заказчика должны быть оснащены следующим ПО:

- операционная система WindowsXP/Vista/7/8;
- браузер Internet - Microsoft Internet Explorer 9.0 или выше;
- офисный пакет приложений Microsoft Office версия 2007 или выше;

Для корректной работы пользователей в Системе рабочие станции должны быть оснащены следующим аппаратными средствами:

- процессор стандартной архитектуры с тактовой частотой не менее 2,33 ГГц или выше;

- объем оперативной памяти – не менее 2 Гб;
- Свободный объем дискового пространства – не менее 20 Гб.

8.1. Лицензионное ПО по распознаванию текста

Поставщик в рамках осуществления работ по оцифровке должен предоставить программное обеспечение оптического распознавания текста и конвертации документов (Recognition), которое должно обладать следующими возможностями:

Распознавание/конвертация текста

Распознавание/конвертация многоязычных документов;

Режим быстрого распознавания/конвертации;

Распознавание различных типов текста: типографская печать, пишущая машинка, матричный принтер, MICR (E13B), OCR-A, OCR-B, текст, напечатанный готическим шрифтом, текст, полученный с факса;

Распознавание цифр;

Возможность подключения пользовательских эталонов для распознавания/конвертации текстов, напечатанных нестандартными шрифтами;

Возможность подключения шаблона области.

Распознавание и обработка специальных символов

Автоматическая идентификация и распознавание штрих-кодов;

Одномерные: Check Code 39, Check Interleaved 2 of 5, Codabar, Code 128, Code 39, Code 39 without asterisk, Code 93, EAN 8, EAN 13, IATA 2 of 5, Industrial 2 of 5, Interleaved 2 of 5, Matrix 2 of 5, Patch Code, PostNet, UCC-128, UPC-A, UPC-E;

Двумерные: Aztec, Data Matrix, PDF 417, QR Code;

Возможность распознавать на изображениях только штрих-коды;

Автоматическая идентификация и распознавание штрих-кодов, расположенных на странице под любым углом к горизонтали.

Предварительная обработка изображений

Автоматическое определение ориентации страницы текста (определение поворота на 0, 90, 180, 270 градусов);

Исправление перекоса изображений;

Очищение изображения от фоновых шумов;

Разделение сдвоенных страниц;

Исправлять разрешение;

Преобразовать цветные изображения и изображения «в сером» в черно-белые;

Возможность удаления пустых страниц.

Контроль качества распознавания

Возможность установки порога качества распознавания текста, основанного на допустимом количестве неуверенно распознанных символов на странице;

Возможность настройки проверки оператором всех распознанных страниц, или проверки только выборочных страниц, распознанных с качеством ниже заданного порога;

Возможность визуальной проверки и корректировка распознанного текста на станции верификации;

Возможность автоматической отмены выполнения задания при достижении установленного временного порога на обработку задания;

Возможность настроить дополнительную обработку изображений и файлов XML-результата для забракованных заданий при помощи скрипта.

Разделение и сборка документов

Разделение документов в потоке по пустым листам, листам с разделительным штрих-кодом, по фиксированному количеству страниц в документе;

Разделение документов с помощью скрипта;

Возможность сборки одностраничных изображений из каждой обрабатываемой подпапки в отдельный документ;

Возможность создавать один документ для каждого файла в задании на обработку;

Удаление пустых страниц-разделителей или страниц со штрих-кодами после разделения;

Возможность разделять документы с помощью скрипта.

Индексирование

Возможность присвоить документу тип и атрибуты;

Распознанный текст переносится в поля атрибутов по клику мыши;

Проверка формата введенных атрибутов по регулярным выражениям;

Возможность использования скрипта, например, для индексирования документов на основе распознанного текста.

Экспорт результатов обработки

Присвоение имени исходящему файлу согласно заданным правилам: по штрих-коду, по дате и времени и т.п.;

Возможность сохранения полного оформления исходного документа (стили, размер и тип шрифтов), а также сохранения картинок в документе-результате конвертирования/распознавания;

Возможность создавать PDF-документы с ультравысоким сжатием MRC (MixedRasterContent), позволяя сохранять высокое качество изображения при этом значительно уменьшать размер PDF-файла документов;

Возможность одновременного сохранения результата распознавания/конвертации в нескольких различных выходных форматах.

Возможность сохранять исходные изображения.

Программное обеспечение должно обеспечивать работу со следующими

форматами данных

Импорт изображений:

- **PDF:**

Файлы в формате PDF (версии 1.7 или более ранней)

- **BMP:**

несжатый черно-белый;

4- и 8-битный — несжатый Palette;
4- и 8-битный — Palette с RLE-сжатием;
16-битный — несжатый, несжатый Mask;
24-битный — несжатый;
32-битный — несжатый, несжатый Mask.

• **TIFF:**

черно-белый — несжатый, CCITT3, CCITT3FAX, CCITT4, Packbits, ZIP, LZW;
серый — несжатый, Packbits, JPEG, ZIP, LZW;
24-битный цветной — несжатый, JPEG, ZIP, LZW;
1-, 4-, 8-битная палитра — несжатый, Packbits, ZIP, LZW.

• **JPEG:**

серый, цветной

• **JPEG 2000:**

серый — Part 1;
цветной — Part 1.

• **JBIG2:**

черно-белый

• **PNG:**

черно-белый, серый, цветной

• **GIF:**

черно-белый, LZW;
2-, 3-, 4-, 5-, 6-, 7-, 8-битная палитра, LZW.

• **DjVu:**

черно-белый, серый, цветной

• **PCX, DCX:**

черно-белый;
2-, 4- и 8-битная палитра;
24-битный цветной.

• **WDP:**

черно-белый, серый, цветной

Экспорт данных:

Microsoft Word (*.DOC, *.DOCX);

Microsoft Excel (*.XLS, *.XLSX);

Rich Text Format (*.RTF);

Microsoft Word XML (*.XML) (только для Microsoft Office Word 2003);

PDF, PDF/A;

EPUB;

HTML;

CSV;

Plain text (*.TXT);

TIFF ;

JPEG, JPEG 2000;

XML со схемой представления данных (*.XML)¹

JBIG2;

Alto XML;

EPUB;

XML со схемой представления данных (*.XML) должен включать:

сохранение атрибутов символов (в том числе, координаты символов);

сохранение дополнительных атрибутов символов (в том числе, был ли символ неуверенно распознан, было ли слово найдено в словаре);

сохранение координат символов относительно исходного изображения;

сохранение форматирования символов (в том числе, размер шрифта, стиль шрифта).

Программное обеспечение должно обеспечивать следующую организацию работы:

Централизованное распознавание/конвертация документов

Многопользовательский режим работы по распознаванию/конвертации документов;

Вся обработка документов должна осуществляться на сервере Поставщика;

Возможность настройки параметров процессов обработки, расписаний, приоритетов выполнения заданий;

Управление очередностью выполнения заданий;

Возможность составления расписания активности обработки (период повторения, время выполнения задания, время действия расписания);

Передача изображений на обработку (расознавание/конвертацию) должна осуществляться с помощью:

стандартных функций обмена файлами, предоставляемыми сетевой операционной системой (FTP, ресурсов общего доступа);

электронной почты (почтовый ящик Exchange, почтовый сервер POP3);

непосредственно с сетевых сканеров и МФУ.

Передача обработанных документов пользователям должна осуществляться с помощью:

стандартных функций обмена файлами, предоставляемыми сетевой операционной системой (FTP, ресурсов общего доступа);

электронной почты (SMTP-сервер, Exchange Server);

путем публикации документов в библиотеки Microsoft Office Share Point Server.

Масштабируемость и отказоустойчивость:

- поддержка многопроцессорных систем: возможность запуска нескольких параллельных процессов по распознаванию/конвертации;

- масштабируемость на несколько компьютеров в локальной сети: возможность подключения дополнительных станций обработки;

- возможность автоматического управления очередью заданий и распределения объема работы между станциями обработки;

- контроль работоспособности Станций обработки Менеджером сервера в ходе выполнения заданий;

- Автоматический запуск системы при запуске Windows.

Интеграция с другими приложениями:

Программное обеспечение должно предоставлять два типа Прикладного программного интерфейса для интеграции с внешними системами:

СОМ-совместимый прикладной программный интерфейс (ComponentObjectModel), возможность создавать приложения в средах разработки C/C++, VisualBasic и любых других, поддерживающих СОМ-компоненты, и языках сценариев;

Прикладной программный интерфейс, предоставляемый через веб-сервис – WebServiceAPI (использование SOAP и http, поддержка работы с удалёнными приложениями, поддержка кросс-платформенности);

Возможность управления настройками обработки заданий из других приложений с помощью XML-описаний, как из почтовой программы, так и непосредственно с сетевых сканеров и МФУ;

Интеграция с внешними системами при помощи скриптов API настройки типов обрабатываемых документов и атрибутов документов;

Интеграция с поисковыми серверами Google Search Appliance и Microsoft Office SharePoint Server;

Поддержка интерфейса Microsoft I Filter для Microsoft Office SharePoint Server и для персональных компьютеров на базе Microsoft Windows.

Администрирование:

Общий интерфейс управления для конфигурации и мониторинга системы;

Детальная настройка, проверка состояния процесса обработки документов и управление списком сценариев обработки;

Просмотр очереди заданий и управление ей;

Управление учетными записями пользователей;

Ведение журнала событий;

Отправка уведомлений по электронной почте администратору о сбоях в работе системы;

Объем используемой лицензии по распознаванию текста должен быть не менее (3 103 000 страниц) заявленного в конкурсной документации;

Потенциальный поставщик должен подтвердить право использования промышленного программного обеспечения.

9. Требования к порядку контроля и приемки выполненных услуг

Услуга считается выполненной после подписания акта выполненных услуг между Заказчиком и Исполнителем.

Цель приемосдаточных испытаний состоит в проверке соответствия ЭКД характеристикам соответствующим требованиям данной технической спецификации.

Приемосдаточные испытания ЭКД организуются Заказчиком, совместно с Исполнителем.

Прием-передача выполненных работ по созданию программного обеспечения и результатов сканирования и ретроконверсии осуществляется путем подписания акта выполненных работ с проведением, при необходимости, процедуры выборочного контроля.

10. Требования к Поставщику услуг

Потенциальный поставщик предоставляет перечень специалистов, которые будут задействованы в выполнении работ по потоковой обработке документов (сканирование), а также загрузке в базу данных системы электронного архива:

Требование к качеству услуг и составу исполнителей	Подтвердить
Не менее 2 специалистов в области архивного дела с опытом работы по этой специальности не менее 3-х лет	Нотариально заверенная копия либо оригинал сертификата/диплома о прохождении обучения и копия трудовой книги
Не менее 1 сертифицированного менеджера проекта уровня PMP/IPMA	Нотариально заверенная копия либо оригинал сертификата
Не менее 4-х специалистов с сертификатами OpenText Certified Capture Center Professional	Нотариально заверенная копия либо оригинал сертификата
Не менее 4 сертифицированных специалистов в области распознавания, верификации и индексирования	Нотариально заверенная копия либо оригинал сертификата
Не менее 1 сертифицированного специалиста по администрированию и настройке промышленного программного обеспечения по распознаванию текста	Нотариально заверенная копия либо оригинал сертификата
Потенциальный поставщик должен подтвердить право использования программного обеспечения по распознаванию текста объемом не менее 3 103 000 страниц	Нотариально заверенная копия либо оригинал документа

Подписи Сторон:

Заказчик:

М.П.

Поставщик:

М.П.

Председатель тендерной комиссии
Управляющий директор
По информационным технологиям



Исмагамбетов Т.О.

