🚯 Sherpa RPA

Оглавление

Шаблонизатор Sherpa RPA2
1.Интерфейс Шаблонизатора3
1.1 Основное меню Шаблонизатора4
1.1.1 Меню «Файл»6
1.1.2 Меню «Обработка»7
1.1.3 Открыть исходный файл9
1.1.4 Обработать документ10
1.1.5 Проверить11
1.1.6 Добавить атрибут12
1.1.7 Выполнить полное сканирование13
1.1.8 Умный помощник15
1.1.9 Масштаб
1.1.10 Меню «Отображать»17
1.2 Центральная рабочая область Шаблонизатора18
1.3 Панель настроек
1.3.1 Настройки
1.3.2 Якоря
1.3.3 Атрибуты
1.3.4 Таблицы
1.3.5 Чертежи таблиц
2. Действия в Sherpa Designer после создания шаблона
2.1 Значения переменных
2.2 Работа с многостраничными документами
3. Примеры
3.1 Создание шаблона на примере документа «Справка о стоимости выполненных работ»
3.1.1 Предобработка документа67
3.1.2 Создание шаблона
3.1.3 Работа с таблицами72
3.1.4 Якоря
3.1.5 Создание атрибутов
3.2 Создание шаблона с фиксированными атрибутами на примере документа «Счет на оплату»95



Шаблонизатор Sherpa RPA

Шаблонизатор представляет собой инструмент создания шаблонов (визуальный редактор) и является неотъемлемой частью Дизайнера Sherpa RPA. Шаблонизатор не привязан к каким-либо облачным сервисам и входит в комплект поставки.

Основные возможности

• Определение типа документа (текстовый, табличный; одно- или многостраничный и т.д.). В случае, когда роботу требуется создать несколько шаблонов по различным типам документа (например, Акт отгрузки товара, Акт оказания услуг), Шаблонизатор автоматически подбирает наиболее подходящий шаблон из списка;

• Исправление перспективы и перекосов скан-копии (фото) обрабатываемого документа;

- Зонирование и сегментация документов;
- Автоматический поиск и разметка таблиц, в том числе без видимых

границ;

- Динамическая привязка содержимого;
- Проверка наличия печатей, подписей, рукописных текстов;
- Извлечение корректной информации из скан-копии/фото

документа.

Обработка сложных документов

Процесс обработки сложных документов выглядит следующим образом:

№ п/п	Этап обработки	Описание
1.	Коррекция и типизация	Определение типов документов, разбор по типам и страницам Коррекция перекосов и перспективы Зонирование и сегментация



2.	Предобработка Шаблонизатором	Большая часть шаблона создаётся автоматически Автоподбор подходящего шаблона Автоопределение границ таблиц
3.	Обработка OCR	Онлайн или офлайн 5 модулей (движков) на выбор Внешние или встроенные в платформу
4.	Постобработка документа	ML & NLP Man in the loop Проверка наличия подписей, печатей, галочек, рукописных текстов
5.	Занесение в ИС	SAP Oracle R12 1С Битрикс 24 и другие

1. Интерфейс Шаблонизатора

Запуск Шаблонизатора осуществляется на верхней панели основного меню Дизайнера во вкладке «Разработка» — раздел «Запись» — «Извлечь по шаблону».

🕨 Отладка 🔹 🔳 Сто 6 Ð Ð € € 모 🖽 Sherpa Explorer D -0-Ŧ **O**(R ? 🔿 Шагвперёд 🕴 Шаг Git Commit • Опублик Помощ Точка останова • зная панель Быстрый Старт × Н Блоки Проект Биб. Добро пожаловать в Sherpa RPA! 📕 Роботы для Windows иска наберите название блока Новый проект0 Новый Открыть Новый проект проект... проект... Новый проект \bigcirc Клонировать проект из <u>Git</u> Windows КрестикиНолики Процес Tестовый пример Calc Example Открыть папку ToDo LOG Проект Calc Example <u>1С ПриходнаяНакладная</u> Print Certificates and Badges проектов Лог \Diamond WhatsappSend $\mathbf{?}$ \$a = 1 Тестовый Присвоить значение переменно Switch/Case ٢ Цикл Для каждого (Список) Цикл Для каждого (Словарь) \Diamond \Diamond (1) (B) Пауза 0 Им Началы Цикл Пока Тип Цикл C# Ş PS> полнить код Python олниті С# выражени PowerShe VB Глобал ۲ Выполнить код Try/Catc Контейне

При нажатии на кнопку «Извлечь по шаблону» открывается обработка «Генератор шаблонов» (далее Шаблонизатор).

Пользовательский интерфейс Шаблонизатора состоит из следующих основных областей:

1) Основное меню (верхняя панель);



2) Центральная рабочая область – область для работы с шаблоном

обрабатываемого документа;

3) Панель настроек.

🐯 New pattern - Генератор шаблонов	– 🗆 X
Файл Обработка	
🗑 🏷 1) Открыть исходный файл 🕨 2) Обработать документ 👻 🖪 Проверить 🕂 3) Добавить атрибут 🔁 4) Выполнить полное сканирование 🋞 Умный помощник	
Image: 100%) Остображени: Опора (Атрибуты) О Царона (С. П. К.	Истройки ТезенасОСК ОСК масшлаб 2 Явык распознавания 2 Явык распознавания 7 Примсе изовления и р 0 ОСК размер нейки пс 40 0 Фильтр поризовлать 15 3 Апотовераневание сту Тлие Апотовераневание сту Тлие Апотовераневание сту Тлие Апотовераневание сту Тлие Обрабатывать вся стр Fade Синемать блоки Обрабатывать вся стр Fade Синемать блоки Обрабатывать вся стр Fade Собрабатывать вся стр Fade Обрабатывать вся стр Fade Собрабатывараспознавание Тессть Табляма распознавание тессть Тобнои превы Процент изображения Водать соторык будат производится распознавание Мадать, которык будат производится распознавание Маскизмой В Пастройки Водать В Пастройки Водать В Пастройки Водать В Пасямы Водаты

1.1 Основное меню Шаблонизатора

Верхняя панель основного меню позволяет увидеть данные, вычисленные встроенным OCR — модулем, а также предлагает основные действия для создания и обработки шаблонов.





Элементы	интерфейса	основного меню	Шаблонизатора.
	11		

№ п/п	Элемент интерфейса	Описание элемента интерфейса
1.	меню «Файл»	Позволяет использовать базовые функции по управлению файлами документов и шаблонов.
2.	меню «Обработка»	Позволяет производить основные действия по обработке документа.
3.	кнопка «Открыть исходный файл»	Позволяет выбрать и открыть ранее сохраненный сканированный документ, который будет использоваться в качестве образца при создании шаблона для распознавания роботом определенного типа документа.
4.	кнопка «Обработать документ»	Позволяет запустить обработку документа по выбранным параметрам для отображения (Якоря, Атрибуты, Таблицы, Чертежи таблиц, ОСК блоки, ОСК таблицы, ОСК линии).
5.	кнопка «Проверить»	Позволяет визуально обновить макет шаблона после обработки документа и внесения каких-либо правок в шаблон (создание/редактирование якоря/атрибута/таблицы) без повторной отправки документа на обработку OCR.
6.	кнопка «Добавить атрибут»	Позволяет определить и обозначить области документа, данные которых будут распознаваться роботом и импортироваться в таблицу «Результаты».
7.	кнопка «Выполнить полное сканирование»	Позволяет выполнить полное сканирование.
8.	кнопка «Умный помощник»	Позволяет проанализировать документ и создать шаблон для документа.



9.	переключатель «Масштаб»	Позволяет задать масштаб при работе со сканированным документом.
10.	Меню «Отображать»	Меню «Отображать» позволяет управлять отображением якорей, атрибутов, таблиц и чертежей таблиц, ОСК блоков, ОСК таблиц, ОСК линий в рабочей области.

1.1.1 Меню «Файл»

Элементы интерфейса основного меню Шаблонизатора. Меню «Файл».

№ п/п	Элемент интерфейса	Описание элемента интерфейса
1.	функция «Новый шаблон»	Позволяет создать новый шаблон.
2.	функция «Открыть исходный файл»	Позволяет выбрать исходный файл для создания шаблона. При нажатии на кнопку открывается окно для поиска исходного файла с возможность указать путь и название файла.
3.	функция «Открыть файл шаблона»	Позволяет выбрать один из ранее созданных шаблон по указанному пути.
4.	функция «Сохранить шаблон»	Позволяет сохранить созданный шаблон по указанному пути. Имя шаблона задается вручную. Также, для сохранения можно воспользоваться сочетанием клавиш Ctrl+S или нажать на в левом верхнем углу панели основного меню. Имя файла нужно указывать, исходя из вида документа, например, «ТОРГ 12». Расширение сохраняемого файла .json
5.	функция «Сохранить как»	Позволяет сохранить созданный шаблон по указанному пути. Имя шаблона задается вручную. Расширение сохраняемого файла .json



6.	функция «Последние файлы»	Позволяет выбрать один из ранее сохраненных файлов из выпадающего списка.
7.	функция «Автоматическое открытие предыдущих файлов»	Позволяет включить автоматического открытие предыдущих файлов. При каждом открытии Шаблонизатора автоматически открывает файл, с которым велась работа до закрытия окна Шаблонизатора. Для активации функции необходимо кликнуть левой клавишей мыши по строке до появления
8.	функция «Выход»	Позволяет выйти из Шаблонизатора.

Φai	iл	Обработка	
₽.	Новый шаблон		
5	Открыть исходный файл		
Э	 Открыть файл шаблона 		
0	Co	хранить шаблон	Ctrl+S
	Co	хранить как	
	П	оследние файлы	•
~	A	томатическое открытие пре	дыдущих файлов
€	Выход		

1.1.2 Меню «Обработка»

Элементы интерфейса основного меню Шаблонизатора. Меню «Обработка».

№ п/п	Элемент интерфейса	Описание элемента интерфейса
1.	функция «Обработать документ»	Позволяет выполнить обработку текущего файла по заданным инструментам и атрибутам, выбранным в верхней панели основного меню (Якоря, Атрибуты, Таблицы, Чертежи таблиц, ОСК-блоки, ОСК-таблицы, ОСК-линии). Также, для обработки документа можно воспользоваться сочетанием клавиш Ctrl+R.

Sherpa RPA

2.	функция «Обработать элементы»	Позволяет выбрать один из ранее созданных шаблонов по указанному пути для его дальнейшей обработки по заданным инструментам и атрибутам, выбранным в верхней панели основного меню. Также, для обработки элементов можно нажать F5.
3.	функция «Выполнить полное сканирование»	Позволяет выполнить полное сканирование. После выполнения действия открывается окно «Результат полного сканирования».
4.	функция «Следующая страница»	Позволяет перейти на следующую страницу в случае, если документ многостраничный. Также, для перехода на следующую страницу можно воспользоваться сочетанием клавиш Ctrl+Right.
5.	функция «Предыдущая страница»	Позволяет перейти на предыдущую страницу в случае, если документ многостраничный. Также, для перехода на следующую страницу можно воспользоваться сочетанием клавиш Ctrl+Left.
6.	функция «Перейти на страницу»	Позволяет перейти на страницу по ее номеру. Также, для перехода на страницу по ее номеру можно воспользоваться сочетанием клавиш Ctrl+G.
7.	функция «Автоматическая обработка документов на следующей странице»	Позволяет выполнить автоматическую обработку следующей открытой страницы по заданным инструментам и атрибутам, выбранным в верхней панели основного меню для текущей страницы. Используется в случае, когда в одном файле находится несколько однотипных одностраничных документов, размещенных на одной странице (например, документ ТОРГ 12). При нажатии кнопки открывается Следующая страница и распознается следующий документ.



	При работе с многостраничным документом, содержащим в себе разные типы документации, не рекомендуем использовать эту функцию, т.к. робот будет распознавать все якоря, созданные на каждой странице документа. Многократный переход между страницами
	может увеличивать время работы робота.



1.1.3 Открыть исходный файл

Шаблонизатор позволяет выбрать и открыть ранее сохраненный сканированный документ, который будет использоваться в качестве образца при создании шаблона для распознавания роботом определенного типа документа.

Файл может быть представлен в одном из форматов: .pdf, .png, .jpeg, .jpg. Файл .pdf может быть многостраничным.



1.1.4 Обработать документ

Кнопка «Обработать документ» позволяет запустить обработку документа по выбранным параметрам для отображения якорей, атрибутов, таблиц, чертежей таблиц, OCR блоков, OCR таблиц, OCR линий в рабочей области. Обработанный документ выглядит как показано на скриншоте.

arrafi: 35% • 🗩 🖉 Orotip	ажаты: 🕢 Якора 🔘 Атриб	уты 👁 Таблицы	Черте	ежи таблиц	OCR 6noxx OCR	таблиции 🕸 OCR Л	terarut	Настройки	
								OCR MO SYTE	YandexVisi
		-	-	0.1				OCR Macuitat	2
		AKT	верки					History pactrophysics and property	
	B3/JetMH	ых расчетов зап	териод 9	Mecallee 2	2023 1			Doctore witholware travets at PD	No
	a Dufin	manager (and incomplete	Con color and	crno"least	title of our in "			ОСЯ размер жейки по горизонта:	NM 70
	ON DEPART	no goroeccy 202	302 of 0	3 02 2022	LAPIN VILLE			ОСП размер инойки по вертикали	70
								Фильтр горизонтальных личий	5
The construction of the second second	THE PROPERTY AND LODGE THE PLANE	PERSONAL PROPERTY AND		Paral Star Intel	Art "Linning" Linners	An and Life Instantion of C	VOIDER PROVIDENT	Антоповорот страницы	True
R Dufference in Canada Contractor	акалонетной общество	Reservence	wate is d	participation .	E apyroit choose	COCTORNED HACTOR	LIGHT DET CRECKS	Автовыраение страница	True
B TOM INTO C OC TOBRANI B SAMARANA DOC	Seton on damain vieta co	Mayoure			a serior stopone	Contraction Production	orthone many respectively.	Критерия дляны линии. Автосырая	Educ
Photosophi de Manager and		and completely	Pla minute	- las multimates	AND DATURDURAD ARTICLES	and "Designation of "		Displate islears for among and	False
in period in manufilling , pyo.			то дани	and HADDING	nun avrikowebieto opifieto	Serviceous		Склемвать блоки	False
Дата Документ	Дебет	Кредит	Дята		Документ	Дебет	Кредит	Разделить блоки	Faise
03.02.23 Dootana (57.05.03.02.2023)	100 103 02		Сальдо	начальное				Игнорировать сцаи5ки вкорей	False
21.02.23 Onnata (865 or 21.02.2023)	106 193,92	108 193.9	2					Удалить блоки прелышающие раз	мер
28.02.23 Продажа (119 or 28.02.2023)	108 193,92							Процент изображения для распоз	Falsara Falsa
16.03.23 Onnets (1431 or 16.05.2023)	100 101 00	108 193.9	2	-		-		nammine+atts	raise
25.04.23 Onvera (2412 or 25.04.2923)	100 193,92	108 193.9	2	-		-	-		
30.04.23 Rpqgawa (242 or 30.04.2023)	108 193,82								
25.05.23 Oninata (3056 of 25.05.2023)	100 100 10	108 193.9	2	_		_			
21.05.23 (Dpogawa (305 dr 31.05.2023) 27.06.23 (Dpogra (3654 cr 27.06.2023)	108 192,92	108 193 9		-					
30.06.23 Ppogawa (369 or 30.06.2023)	100 193,92	100 (102.8	-						
25.07.23 Onnata (4220 of 25.07.2023)		108 193,9	2						
31,07,23 (1pqgawa (435 or 31,07,2023) 29,08,23 (2pqara (4969 or 29,08,2023)	106 193,92	104 193 9	2	-		-	-		
51.08.23 Ppogawa (500 or 31.08.2023)	108 193,92	100.190,8		-					
28.09.23 Onneta (5462 or 28.09.2023)		108 193,9	2						
30.09.23 Opgawa (561 et 30.09.2023)	108 193,92	865 551 3	E Oforore	A DE DROMOS		-	+		
Сальдо вонечное	108 193,92		Canado	ALL HER YORGE					
Падариений АОСЧалина? ин 19.55.2022 задотнотичеств флости у АО 19.65.302 у убл. ССП в посемы такоты сти дан ОСГАСТивлика? (Завес тектовь токурального деректора по зис	C"Manesa) annoctu tpik py5/rel 92 koneikul provant a darnaschik	D	100 garma	nerod Galurond	а алиулонеціной общаство. "За приой общаство «Зампаника»	All Bryon, B		ОСЯ размер менйки по пертикал	•
-	(Maanoe H.M.)		normal la						
MIL			MIL						
								Hac gover	
								A Skepa	
								Т Атрибуты	
								Tativiai	



Также, можно выбрать один или несколько критериев на верхней панели основного меню Шаблонизатора.



При нажатии на кнопку появляется возможность произвести обработку следующей страницы (Ctrl+Right) или предыдущей страницы (Ctrl+Left), перейти на страницу (Ctrl+G). Кнопка «Открыть страницу повторно» позволяет открыть исходный документ без внесенных изменений в настройках Шаблонизатора. Данная кнопка используется для проверки корректности заданных настроек в Шаблонизаторе. Таким образом можно проверить, например, срабатывает ли настройка «Автоповорот страницы». Если при нажатии кнопки «Открыть страницу повторно» мы видим, что документ изначально был неверно повернут, значит настройки Шаблонизатора были заданы верно и при обработке сработал Автоповорот страницы.

2) Обработать документ	•
>	Следующая страница	Ctrl+Right
<	Предыдущая страница	Ctrl+Left
Г.	Перейти на страницу	Ctrl+G
С	Открыть страницу повт	орно

1.1.5 Проверить

Кнопка «Проверить» позволяет проверить корректность отображения табличной части, из которой будут импортироваться данные. Также проверку можно запустить нажатием клавиши F5.

🐼 Sherpa RPA

		A DECEMBER OF A		m. 1 + 3) /lofation and	nefor >0 Re		Autor of Autor and Autor and	and the second se			
				N Same Street		0	antan SOCREAN SOCRE	time to oce a		Tofamore	
<form></form>	<form></form>	• Macu	1202 35% • 20 0100pascare	с фукора фукрабут	ai 💽 taquividar	CO HEPTERO	и таолиц фосколоки фоскта	ounder & OCK 1	neara		
		1 2	Мы, некеподписавляеся, Заместитель иПубличное видоо	взанины и Публич по генерального дарект нерное общество "Зе	Акт с х расчетов за п между АО ное акционерно з договору 2023 тора по эконом эмпянка"	верки ериод 9 м "Малина" е общести 02 от 03.0 иже и фи	есяцев 2023 г. 10 "Земланика" 22 2022 нансам АО "Малина" Иванов И с другой стороны,	зан Иванович, с составили насто	одной стороны, иднай акт сверки	Total and the second seco	1 K. F 14000
Image:	Image:		To assess AO 'Manua' ord	The garman y were crieg	divertee.	file name	и публиканого акционелного общества	"Jen#19/04/3"	1 1	Townee разениство изтонок False	
 	Image: Instantion Image: I		the first of the second states			I TO MANA	a openinento o antipioneginori o congectas	Jesecona		Колонки на первой странице — False	
Image: Sing Sing Sing Sing Sing Sing Sing Sing	A Discrete diversity of the discrete dive		Дата Документ Сальдо начальное	(Jaces	Rpager	Сальдо на	DORVMENT	Defer .	Rptdel?	Оцибка если таблица пустан False	
Image: Distribution of the cost of	Image: Normal of the state		03.02.23 Deosewa (57 or 03.02.2023)	108 193,92						> Колонеи Колонеи (8)	
Image: Normal Normate Normal Normal Normal Normal Normal Norma	Image: Normal Normate Normal Normal Normal Normal Normal Norma		21.02.23 Onnana (665 of 21.02.2023) 28.02.23 Destance (119.of 28.02.2023)	100 103 02	108 193,92						
Image: Section 1 = 1 = 1 = 1 = 1 = 1 = 1 = 1 = 1 = 1	1 1		16.03.23 Onnava (1431 or 16.03.2023)		108 193.92						
	No. Image: No. 100 (100 (100 (100 (100 (100 (100 (100		31.03.23 Продана (176 от 31.03.2023)	108 193,92							
Image: Section 2013 (Section 2013) Image: Section 2014 (Section 2013) Image: Section 2014 (Section 2014) Image: Section	Image: Normal Context (Section 1990) Image: Normal Context (Section 1990) Image: Normal Context (Section 1990) Image: Normal Context (Section 1990) Image: Normal Context (Section 1990) Image: Normal Context (Section 1990) Image: Normal Context (Section 1990) Image: Normal Context (Section 1990) Image: Normal Context (Section 1990) Image: Normal Context (Section 1990) Image: Normal Context (Section 1990) Image: Normal Context (Section 1990) Image: Normal Context (Section 1990) Image: Normal Context (Section 1990) Image: Normal Context (Section 1990) Image: Normal Context (Section 1990) Image: Normal Context (Section 1990) Image: Normal Context (Section 1990) Image: Normal Context (Section 1990) Image: Normal Context (Section 1990) Image: Normal Context (Section 1990) Image: Normal Context (Section 1990) Image: Normal Context (Section 1990) Image: Normal Context (Section 1990) Image: Normal Context (Section 1990) Image: Normal Context (Section 1990) Image: Normal Context (Section 1990) Image: Normal Context (Section 1990) Image: Normal Context (Section 1990) Image: Normal Context (Section 1990) Image: Normal Context (Section 1990) Image: Normal Context (Section 1990) Image: Normal Context (Section 1990)		30.04.25 Ppogama (242 or 30.04.2023)	108 193,92	100 100.04						
Image: State Stat	N/2 0 00.000 <td< td=""><td></td><td>25.05.23 Onnata (3056 or 25.05.2023)</td><td></td><td>108 193.92</td><td></td><td></td><td></td><td></td><td></td><td></td></td<>		25.05.23 Onnata (3056 or 25.05.2023)		108 193.92						
Image:	1021 Transmit Vision 104 100 </td <td></td> <td>27.06.23 Centatia (208.07.27.06.2023)</td> <td>106.193,92</td> <td>108 193 92</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td>		27.06.23 Centatia (208.07.27.06.2023)	106.193,92	108 193 92						
Image: Section 100 (String 100 (Str	Image: Distance of the state of the sta		30.06.23 Opogena (309 or 30.06.2023)	108 193,92							
Image: State Stat	101321 10112101111111 10112101111111111111111111111111111111		25.07.23 Onviana (4220 or 25.07 2023)	300 893 82	108 193,92						
133 Depares 00 100 2033 00 100 203<	Image: Display and State of 100 2001 image: Display and State of 100 2001 <td></td> <td>29.58.23 Cenata (4969 or 29.58 2023)</td> <td>100 102,02</td> <td>108 193.92</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td>		29.58.23 Cenata (4969 or 29.58 2023)	100 102,02	108 193.92						
Image: Strategie (Strategie (Strate	1 100.100.00 <t< td=""><td></td><td>31.06.23 Opigawa (500 or 31.08.2023)</td><td>108 193,92</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></t<>		31.06.23 Opigawa (500 or 31.08.2023)	108 193,92							
Objective in transmit 277 140 ml 481 053 20 000000 mit membra Objective in transmit 181 053 20 000000 mit membra 181 053 20 000000 mit membra Programmed u0 20 0000000 mit membra Dispenses (Pydenove superspects dispetts) "Destruct" 181 052 pid (250 membra) Programmed u0 20 0000000 mit membra Dispenses (Pydenove superspects dispetts) "Destruct" 181 052 pid (250 membra) Programmed u0 20 0000000 mit membra Dispenses (Pydenove superspects dispetts) "Destruct" 181 052 pid (250 membra) Programmed u0 20 00000000 mit membra Dispetts) Programmed under superspects dispetts "Destruct" 181 052 pid (250 membra) Programmed u0 20 00000000000000000000000000000000	Distance in regional 0.27 14/28 Objective in meeting Channel of Visconer ¹ Fill Status of Name Programme Status of Visconer ¹ Fill Status of Name Status of Visconer ¹ Fill Status of Name Status of Visconer ¹ Fill Status of Name Status of Name Ort Optionent augusupport dupits "Statuscone" Mill Mill		30.09.23 Optigasia (561 et 30.09.2023)	108 193,92	108.193.92						
Image: charge charges 100 00.021 Franking charges Image: charge charges Image: charge charges Image: charge charges Image: charge charge Image: charge charge charge Image: charge	Image: Status definitions 100 30.31 Creating definitions Image: Status definitions Programme definitions Programme definitions Programme definitions Image: Status definitions Programme definitions Or Definitions and programme definitions Or Definitions and programme definitions Image: Status definitions Definitions Definitions Definitions Or Definitions and programme definitions Image: Status definitions Image: Status definitions Mathematic definitions Mathematic definitions Mathematic definitions Mathematic definitions Image: Status definitions Image: Status definitions Mathematic definitions Mathematic definitions Mathematic definitions Mathematic definitions Image: Status definitions Image: Status definitions Mathematic definitions Mathematic definitions Mathematic definitions Mathematic definitions Image: Status definitions Mathematic definitions Mathematic definitions Mathematic definitions Image: Status definitions Image: Status definitions Mathematic definitions Mathematic definitions Image: Status definitions Image: Status definitions Image: Status definitions Image: Status definitions Image: Status definitions Im		Обороты за переод	973 745,25	865 551,35	Обороты н	а период				
h gansa fridowni saganspon digens Transmar" h gansa fridowni saganspon digens Transmar" h 1930 fridowni saga	In general Of Vitery ¹ is 19,329 pd, for some states or generating up offer to Teaching ¹ In februe an auguspaper duget to "Teaching ¹ Inference Or AD Vitery ¹ is 19,329 pd, for some states or generating up offer to Teaching ¹ Or Inference auguspaper duget to "Teaching ¹ Inference Description Or Inference auguspaper duget to "Teaching ¹ Or Inference auguspaper duget to "Teaching ¹ Inference Description (Mancie KR)		Сальдо конеченое	108 193,92		Сальдо ко	40-404.54				
Энистопах педералат уфински	Jamman Samana Jamman		По динные АО "Идлина" на 30.05.2023 задолженность в польсу АО "Маля 108 193.32 руб. (Сто восемь тысяч сто деявност От АО "Идлина"	ина" о три рубля 12 колейия).		По данным От Публичен	Публичное акционерное общество "Эвнг ое акционерное общество "Эвнглиния"	theory g		Tadionala	
M. M. M. N. M. M. M. M. M. M. T. Applying T. Ap	Officient 0.000 Officient 0.000 Officient 0.000 Officient 0.000 MIL MIL MIL Image: Control of the control		Заместитеть тенерального директора по экономия	и финански			-			Takinka o monomin or	
NA. NA. Property States of the second secon	MA. MA. Increasion			(Neiamoe N.N.)			L				
Prope Prope Prope Prope Prope Prope Prope Prope Prope Propertment Propertment Propertment	Hampshar Hampshar Hampshar Hampshar Hampshar Totage Totage Hampshar Hampshar Hampshar Hampshar Hampshar Hampshar		M.D.			м.п.					
Interprint	Harpohe H										
A Prope T Annotan T Annotan T Totalan T Totalan T Totalan T Totalan T Totalan T	Anger T Angelyne T Angelyne Totrage Tetrage Vegetrere totfass Vegetrere totfass									C Hactpolior	
T Appendix T Appe	Augusta Augusta									1	
T Analyu Totay	T Andrinu T Andrinu Titinas Venturi talau Venturi talau									sis 780ps	
Teloney-	E Tolinas September tologia September tologia										
🖉 Vegruns rafası	- Veptron tallou									Т Альнбулы	
List table to the second s										T Arputynu Tafonus	
										T Arpadynu Tatovuju Tatovuju	

1.1.6 Добавить атрибут

Атрибуты позволяют определить и обозначить области документа, данные из которых будут распознаваться роботом и импортироваться в таблицу «Результаты».

Существует несколько способов добавления Атрибута:

1) Нажатием кнопки на верхней панели Основного меню.

арыть исходный фай									
зыть исходный фай			-						
	и 2) Обработать документ • • Прове	фить + 3) Добавить :	прибут 🚔 4) Вы	олнить по	пное сканирование (В) Умный пол	ющник			
Magun	таб: 35% - Ø Ø Отоблажа	De O Berra O Arrowf	Martin Co Tatinguese	Q Henrew	a tañasu 🖏 OCR 6.0008 🖏 OCR ta	Annual KO OCR /h	enter	Атрибуты	
1.000 Land		the based of a second and							
								+ T + 41 II / Ipusaska	S S T Y
			A 107 0	a num				Arpedyr :	Без имени
			ANTC	зерки				Има	bed stretest
		E33MMP	ых расчетов за п	pwog 9 M	есяцев 2023 г.			Якорные привязки	Приянски (0)
			между АО	"Малина"				Ограничение личнями	None
		и Публі	чное акционерис	е общести	ю "Земляника"			Корректировать усол	False
			по договору 2023	J2 01 03.0	2.2022			Привнака области	
								Ofinacte	0.0.0
	Мы нижеводнисавшиеся Заместител	ь генерального лисе	ctopa no akonom	же и фи	нансам АО "Малина" Иванов Ив	ан Иванович, с	одной стороны.	Ман количество синиволов	0
	и Публичное ака	онерное общество "	leansea"		C DIVIDA CTODORNA	осталили настоя	LINK AKT CRECKS	Terr dareas a	Test
	B TOM, NTO COCTORISIE BRANNING RACINETO	on no dalella volita co	envouee'					PopMat snakestra	
	the search of the second	of the parents of the tarter	etthroates.					Одибка если значение пустое	False
	По данных АО "Малина", руб.			По данных	а публичного акционерного общества	Земпяника		Проверять экорные привязки	False
			Krieturr	Data	Conversion	Defer	Kontor?	Режим работы	Default
	Data Document	Deter				and an and a state of the state		Onicalise	
	Дата Документ Сальдо и мальное	Deter		Сальдо нач	AAAAAAA				
	Дата Документ Сальдо на пальное 03.02.23. Проданна (57 от 03.02.2023)	Дебет 108 193,92		Cambao H	(07546)¢				
	Дата Дохумент Сальдо на чальное 01.02.23 Продания (157 от 03.02.2023) 21.02.23 Октанта (165 от 21.02.2023)	Дебет 108 193.92	108 193,92	Сальдо нач	ianumise				
	Дата Долимент Сальдо на пальное 03.05.23. Продежа (57 от 03.03.2023) 27.02.23. Осната (665 от 21.02.2023) 28.02.23. Продежа (119 от 28.02.2023)	Deber 108 193.92 106 193.62	108 193,92	Сальдо нач	anunit				
	Aarta Approximation Compage in second	Deter 108 193,92 108 193,92	108 193,92	Сальдо нач	6034404				

2) Нажатием кнопки + на правой рабочей панели меню во вкладке «Атрибуты».





1.1.7 Выполнить полное сканирование

Кнопка «Выполнить полное сканирование» позволяет выполнить полное сканирование документа. После нажатия на кнопку открывается окно «Результат полного сканирования», в котором отображаются распознанные данные из документа в виде заполненной таблицы. Если при проверке обнаружены неточности, то их, в любой момент, можно исправить.

Файл	Обработка										
6	1) Открыть исходный файл	2)	06p	аботать докум	ент 🔹 🕨	Проверить	🕂 3) Добави	ть атрибут	🔁 4) Выполнить	полное сканирование	🛞 Умный помощник
Ma	асштаб: 50%	• ,0	ø	Отображать:	• Якоря	 Атрибуты 	 Таблицы 	• Чертеж	ки таблиц 🗞 ОСК	блоки 🛇 ОСК таблиц	ы 🗞 ОСК Линии

1.1.7.1 Результат полного сканирования

Результат полного сканирования представляет собой окно с результатами, полученными в процессе работы робота со сканированным документом по созданному шаблону. В появившейся форме отображаются все атрибуты и таблицы, обнаруженные в документе.

existing bag	Оты Простой режни - О Обновит	T Tose	ри всех окон					
Агрибуты		Tatireue	0					
ther	Surveyor	Ho-	Намменование посковых контона.com			Barrio Sweetenax		
-	Гентодрядник) 000 г Москва Трофинова 8(49) 4912131	Mepho	этапка объектов, акоса выполненных работ.	Koa	Lin-a	py6.a	Crownooth	Коннество
Инвестор	ОООВиктории Москва Краснопрологарская	00:	оторудовании, затрат	-		1101210 (023	1202.000	-
Намер	14		денкантажные работы по сноку падсобных помещения	-	100 000	100 000	100 000	-
Date	31.06.2016		иситах межномнатных парагородон		60 000	60.000	60 000	
enter o	Construction of the second sec		общая станность стройнатерналов		120 000	120 000	120 000	
			1					_
						Viterrar	250 000	
						Сумна НДС	50.400	
					Scero	c yvenow HDC	330 400	
		1.0		10				And and a state of the

Окно с результатами поделено на две табличные части:

 Атрибуты (слева) – в данной табличной части отображаются данные, полученные роботом по заданным Атрибутам и Якорям. Чаще всего, это общая информация о Заказчике, Подрядчике, Инвесторе, например, наименование и реквизиты. Также в Атрибутах указываются Номер и Дата документа.

🐼 Sherpa RPA

 Таблица (справа) – в табличной части отображаются данные, содержащиеся в таблицах сканированного документа, полученные роботом в результате работы по созданному шаблону при помощи ОСR-модулей.

Эти данные могут быть числового формата, например, цена, стоимость, сумма и т.д.

По умолчанию указан режим «Простое сканирование».



При работе с многостраничными документами (если файл содержит в себе два и более документа), на одной из страниц которого указан якорь «LastPage», необходимо использовать «Режим полного распознавания».

🛂 Результат по	лного сканирования	
Режим работы:	Режим полного распознавания	٠

При этом будет отображаться выходной параметр – количество страниц, которые участвовали в работе Шаблонизатора.

Режим работы:	Режим полного распознавания	-
Атрибуты	/	
Имя 🖊	Значение	
CountPages	1	

Полное распознавание будет проводиться с текущей страницы до момента, пока не будет найден якорь «LastPage».

В случае, если мы запускаем «Полное сканирование», находясь не на первой странице многостраничного документа, и, если на одной из страниц указан якорь «LastPage», Шаблонизатор выдаст окно с запросом: «Вы хотите выполнить распознавание с первой страницы?».

🐼 Sherpa RPA

При нажатии кнопки «Да» — алгоритм Шаблонизатора будет распознавать документ с первой страницы документа до момента, пока не будет найден якорь «LastPage».

При нажатии кнопки «Нет» — алгоритм Шаблонизатора будет распознавать документ с текущей страницы документа до момента, пока не будет найден якорь «LastPage».

Macunat: 47%	ogeniā ģaža 🕨 2) Olgadoraru gonguner → [- Ø Ø Orotpanaru: © facça ©	≥ Провереть + 31 Добленть из Аграфуты © Tatanaus © Чарто	pellyr 👌 👌 Demonstra fizikos cale ne raforeg © OCA tonne © OCA raf	esposaeee Servaar SQ-COL/beeser	Berga	
	war and physical densities	Participan		James And	+ - 12 1 Принна 8 8 + +	
Decement op Jamma in span Temp strategy Jamma date Jamma date Department Jamma date Jamma date		Being With Yook and House	анаруулаан алаан алаа Элий талаан алаан алаа алаан биликин алаан а	Alterethi. Annual Transmission (age of the second	Auge Auge	

1.1.8 Умный помощник

Кнопка «Умный помощник» позволяет проанализировать документ и создать шаблон для документа.

Файл	Обработка	
6	 Открыть исходный файл 	🕨 2) Обработать документ 🔹 🕞 Проверить 🛛 🕂 3) Добавить атрибут 🛛 🔁 4) Выполнить полное сканирование 🏼 🎯 Умный помощник
• N	Ласштаб: 50%	🔸 🔎 🔎 Отображать: 💿 Якоря 💿 Атрибуты 💿 Таблицы 💿 Чертежи таблиц 🗞 ОСК блоки 🗞 ОСК таблицы 🗞 ОСК Линии

При нажатии на кнопку «Умный помощник» доступны следующие

настройки.

Элементы интерфейса основного меню Шаблонизатора. Умный помощник. Настройки.

№ п/п	Элемент интерфейса	Описание элемента интерфейса
1.	настройка «Извлечь атрибуты автоматически»	Позволяет роботу определить значения атрибутов автоматически, без прямого указания.

🚯 Sherpa RPA

2.	настройка «Извлечь	Позволяет задать определенные атрибуты
	только указанные	для извлечения их из документа
	атрибуты»	(например, «покупатель» или «продавец»).
3.	Настройка «Извлечь	Позволяет указать прямо атрибуты,
	указанные и	необходимые для извлечения, а также
	автоматические	позволить роботу самому определить поля,
	определенные	которые присутствуют в анализируемом
	атрибуты»	документе.
4.	флаг «Также извлекать	Позволяет извлечь заголовки таблицы из
	таблицы»	документа. Если флаг не стоит, то таблица
		в документе рассматриваться не будет.
5.	флаг «Очистить	Позволяет при каждом запуске очистить в
	шаблон»	шаблоне Якоря, Атрибуты, Чертежи,
		Чертежи таблиц.

🔱 Умный помощник	×
"Умный помощник" проанализирует и автоматически создаст умный шаблон для текущей страницы документа.	
Вы можете написать названия атрибутов, которые "Умный помощник" должен извлечь из документа. Также "Умный помощник" может извлечь названия полей и документа автоматически.	13
🖲 Извлечь атрибуты автоматически 🖂 Также извлекать таблиць	с I
О Извлечь только указанные атрибуты	
О Извлечь указанные и автоматически определенные атрибуты	
Атрибуты для извлечения:	
	,
Этот процесс может занять некоторое время. Вы готовы начать?	
✓ Очистить шаблон Да Нет	

Умный помощник поможет в работе с Шаблонизатором: он покажет, как анализируется документ и создается шаблон.



1.1.9 Масштаб

Переключатель «Масштаб» позволяет задать масштаб при работе со сканированным документом. Размер изображения может быть задан несколькими способами:

- Выбрать значение из выпадающего списка;
- Указать значение вручную;
- Увеличить или уменьшить масштаб при помощи лупы.



1.1.10 Меню «Отображать»

Меню «Отображать» позволяет управлять отображением якорей, атрибутов, таблиц и чертежей таблиц, ОСК блоков, ОСК таблиц, ОСК линий в рабочей области.

Файл	Обработка										
6 ►	1) Открыть исходный фай	л 🕨 2) Об	работать доку	мент 🗕 🕨	Проверить	🕂 3) Добави	ть атрибут 👌 4)	Выполнить полно	е сканирование	🛞 Умный помоц	цник
• Ma	асштаб: 50%	• @ Ø	Отображать:	• Якоря	• Атрибуты	👁 Таблицы	⊙ Чертежи табли	ц 💿 ОСК блоки	OCR таблицы	○ ОСК Линии	

Элементы интерфейса основного меню Шаблонизатора. Меню «Отображать».

№ п/п	Элемент интерфейса	Описание элемента интерфейса
1.	Якоря	Отображает точки на сканированном
		документе, с помощью которых робот
		определяет, в какой части документа
		размещаются определенные данные.
		Якорей может быть сколь угодно много, в
		зависимости от типа документа.

🐼 Sherpa RPA

2.	Атрибуты	Отображает области документа, данные из которых будут распознаваться роботом и импортироваться в таблицу «Результаты».
3.	Таблицы	Отображает границы таблиц, данные из которых будут распознаваться роботом и импортироваться в таблицу «Результаты».
4.	Чертежи таблиц	Отображает чертежи таблиц.
5.	OCR блоки	Отображает блоки, распознанные OCR- модулем.
6.	OCR таблицы	Отображает таблицы, распознанные OCR- модулем.
7.	OCR линии	Отображает линии, распознанные OCR- модулем.

Кнопки Якоря, Атрибуты, Таблицы и Чертежи таблиц относятся напрямую к шаблону. Данные, отображаемые по ним, являются вычисленными после обработки шаблона.

Фа	іл Обработка						
6	늘 1) Открыть исходный ф	айл 🕨 2) Обработать док	умент 👻 🕨 Проверить	🕂 3) Добавить атрибут	🔁 4) Выполнить полно	е сканирование	🛞 Умный помощник
•	Масштаб: 50%	• 🔎 🔎 Отображати	. 💿 Якоря 💿 Атрибуты	👁 Таблицы 💿 Чертежи	и таблиц 💿 ОСК блоки	OCR таблицы	OCR Линии

Кнопки ОСК блоки, ОСК таблицы и ОСК линии используются перед обработкой шаблона для настройки шаблона. Они отображают данные, полученные после обработки документа ОСК модулем.

```
Файл Обработка

        Файл
        Обработка

        Проверить
        + 3) Добавить атрибут

        Анаситаб:
        50%

        Э
        Обработать документ

        Э
        Масштаб:

        50%
        Э

        Э
        Остравание

        Э
        Остравание

        Э
        Остравание

        Э
        Остравание

        Э
        Э

        Э
        Э

        Э
        Э

        Э
        Э

        Э
        Э

        Э
        Э

        Э
        Э

        Э
        Э

        Э
        Э

        Э
        Э

        Э
        Э

        Э
        Э

        Э
        Э
```

1.2 Центральная рабочая область Шаблонизатора

Область для работы с шаблоном обрабатываемого документа. Здесь

отображается исходный файл/сканированный документ, по которому создается шаблон для дальнейшей работы робота.





При наведении курсора на нижней панели отображаются координаты страницы, курсора и области. Это удобно, когда нужно проверить или скорректировать координаты какого-либо атрибута.

В центральной рабочей области отображаются результаты всех манипуляций, которые мы производим на верхней панели основного меню и на боковой панели настроек Шаблонизатора.

1.3 Панель настроек

В правой панели основного меню Шаблонизатора находятся 5 основных блоков для создания и настройки шаблона:

- Настройки;
- Якоря;
- Атрибуты;
- Таблицы;
- Чертежи таблиц.



стите дуле	TesseractOCR
DCR масштаб	2
Азык распознавания	
Азык распознавания для якорей	
Трямое извлечение текста из PDF	No
ОСR размер ячейки по горизонтали	40
CR размер ячейки по вертикали	20
Рильтр горизонтальных линий	5
втоповорот страницы	True
втовыравнивание страницы	True
(ритерий длины линии. Автовыравние	вани 10
Обрабатывать Pdf аннотации	False
Обрабатывать все страницы	False
Склеивать блоки	False
азделить блоки	False
Ігнорировать ошибки якорей	False
далить блоки превышающие размер	0
Іроцент изображения для распознав	ания
айти печать	False
К модуль иль, которым будет производится рас льным механизмом	спознавание текста. Таблицы распознаются
К модуль иль, которым будет производится рас льным механизмом	спознавание текста. Таблицы распознаются
К модуль иль, которым будет производится рас- льным механизмом Настройки	спознавание текста. Таблицы распознаются
Кодуль ль. которым будет производится рас льным механизмом Настройки Якоря	спознавание текста. Таблицы распознаются
К модуль ль, которым будет производится рас льным механизмом Настройки Якоря Атрибуты	спознавание текста. Таблицы распознаются

1.3.1 Настройки

Шаблонизатор использует несколько видов модулей для извлечения данных из исходных файлов.

ОС модуль

ОСR модуль – модуль, которым будет производиться распознавание текста.

В платформу Sherpa RPA интегрированы несколько ОСR-модулей. Два из них поставляются вместе с роботом: Tesseract OCR и Microsoft OCR. Данные модули могут работать офлайн, без подключения к интернету).

Tesseract OCR — это движок оптического распознавания символов (OCR) с открытым исходным кодом, является самой популярной и качественной OCRбиблиотекой. OCR использует нейронные сети для поиска и распознавания текста на изображениях.

Yandex Vision и ABBYY OCR – онлайн-модули, которые используют функции соответствующих облачных сервисов.



Yandex Handwriting — модуль, позволяющий распознавать рукописный текст.

Microsoft OCR — модуль, позволяющий распознавать текст на изображениях и сканированных документах с помощью оптического распознавания символов (OCR).

OCR Space — модуль, позволяющий распознавать кириллический шрифт на изображениях формата .jpg.

ABBYY FineReader – это коммерческий оффлайн-модуль, использование которого требует отдельной лицензии.

Платформа Sherpa RPA позволяет настроить работу сценария с распознаванием изображений и переключаться между этими OCR-модулями в любой момент времени.

Hac	тройки		
	OCR модуль	TesseractOCR	\sim
	OCR масштаб	TesseractOCR	
	Язык распознавания	YandexVision	
	Язык распознавания д	YandexHandwriting	
	Прямое извлечение тек	MicrosoftOCR	
	ОСR размер ячейки по	OCRSpace	
	ОСR размер ячейки по	AbbyyOCR	
	Фильтр горизонтальны	FineReader	

ОСК масштаб

ОСК масштаб — параметр, позволяющий улучшить качество распознавания, если документы имеют плохое разрешение.

В случае высокого и среднего качества сканированного документа рекомендуется оставить значение масштаба равное «2».

Настройки	
OCR модуль	TesseractOCR
OCR масштаб	2

🐼 Sherpa RPA

Язык распознавания – параметр, позволяющий увеличить точность распознавания документа. Также, можно задать несколько языков, использовав в качестве разделителя запятые.

При выборе настройки «Язык распознавания» открывается выпадающий список, в нем с помощью флагов можно отметить нужные для распознавания языки.





Язык распознавания для якорей

🐼 Sherpa RPA

Язык распознавания – параметр, позволяющий указывать конкретный язык распознавания для якорей. Также, можно задать несколько языков, использовав в качестве разделителя запятые.





Прямое извлечение текста из PDF

Прямое извлечение текста из PDF – параметр, позволяющий управлять прямым извлечением текста со страницы.

Возможные значения:

- No отключено;
- Yes используется только прямое извлечение текста;



• Auto - автоматический режим (если на странице нет текста, то будет

реализовано распознавание текста с заданным ОСК модулем).

Настройки		
OCR модуль	TesseractOCR	
OCR масштаб	2	
Язык распознавания		
Язык распознавания для якорей		
Прямое извлечение текста из PDF	No	\sim
ОСЯ размер ячейки по горизонтал	I No	
ОСЯ размер ячейки по вертикали	Yes	
Фильтр горизонтальных линий	Auto	
	INIE	

ОС**R** размер ячейки по горизонтали

ОСП размер ячейки по горизонтали – параметр, позволяющий указать делитель по горизонтали, по которому определяются ячейки таблиц на странице. Значение должно быть больше или равно 1.

Размер ячейки определяется как размер изображения, деленное на это значение.

Шаблонизатор использует два параметра (по горизонтали и по вертикали) для более точного определения таблиц.

астройки	
OCR модуль	TesseractOCR
OCR масштаб	2
Язык распознавания	
Язык распознавания для якорей	
Прямое извлечение текста из PDF	No
ОСЯ размер ячейки по горизонтали	40
ОСR размер ячейки по вертикали	20

По умолчанию заданы параметры:

40 - OCR размер ячейки по горизонтали

Это значение является оптимальным при распознавании документов со стандартными (или близкими к стандартным) размерами ячеек таблиц.



Рекомендуем оставлять эти значения без изменений и менять их только в случае некорректного распознавания (после проверки).

Если при заданных параметрах (значение по горизонтали 40) таблица не будет распознана, то алгоритм Шаблонизатора увеличит это значение на 10 автоматически и осуществит попытку повторно.

ОСК размер ячейки по вертикали

ОСП размер ячейки по вертикали – параметр, позволяющий указать делитель по вертикали, по которому определяются ячейки таблиц на странице. Значение должно быть больше или равно 1.

Размер ячейки определяется как размер изображения, деленное на это значение.

Шаблонизатор использует два параметра (по горизонтали и по вертикали) для более точного определения таблиц.

астройки	
OCR модуль	TesseractOCR
OCR масштаб	2
Язык распознавания	
Язык распознавания для якорей	
Прямое извлечение текста из PDF	No
OCR размер ячейки по горизонтали	40
ОСЯ размер ячейки по вертикали	20

По умолчанию заданы параметры:

20 – OCR размер ячейки по вертикали.

Это значение является оптимальным при распознавании документов со стандартными (или близкими к стандартным) размерами ячеек таблиц.

Рекомендуем оставлять эти значения без изменений и менять их только в случае некорректного распознавания (после проверки).



Если при заданных параметрах (значение по вертикали 20) таблица не будет распознана, то алгоритм Шаблонизатора увеличит это значение на 10 автоматически и осуществит попытку повторно.

Фильтр горизонтальных линий

Фильтр горизонтальных линий – параметр, указывающий процент, меньше которого линии будут проигнорированы. Определяется как отношение горизонтальной линии к ширине изображения. Данный параметр используется для атрибутов.

астройки	
OCR модуль	TesseractOCR
OCR масштаб	2
Язык распознавания	
Язык распознавания для якорей	
Прямое извлечение текста из PDF	No
OCR размер ячейки по горизонтали	40
ОСR размер ячейки по вертикали	20
Фильтр горизонтальных линий	5
	_

По умолчанию значение данного параметра равно 5.

Часто в сканированных документах присутствуют линии, не являющиеся границами таблиц (различные артефакты, которые необходимо отфильтровать). При заданном параметре все лишние артефакты (линии, нарисованные ручкой или линии, появившиеся при некачественном сканировании) будут проигнорированы.

Автоповорот страницы

Автоматический поворот страницы – параметр, позволяющий автоматически разворачивать страницы на угол, кратный 90 градусам (90, 180 и 270). Автоповорот страницы не разворачивает документ на малые углы. По умолчанию значение параметра «True», рекомендуем оставлять его без изменения.

Sherpa RPA

Настройки	
OCR модуль	TesseractOCR
OCR масштаб	2
Язык распознавания	
Язык распознавания для якорей	
Прямое извлечение текста из PDF	No
OCR размер ячейки по горизонтали	40
ОСR размер ячейки по вертикали	20
Фильтр горизонтальных линий	5
Автоповорот страницы	True 🗸
Автовыравнивание страницы	True
Критерий длины линии. Автовыравниван	False
Обрабатывать Pdf аннотации	raise



Автовыравнивание страницы

Автоматическое выравнивание страницы – параметр, позволяющий выравнивать содержимое страницы при некорректном сканировании листа. В отличие от Автоповорота страницы, Автовыравнивание разворачивает документ на малые углы.

тройки	
OCR модуль	TesseractOCR
OCR масштаб	2
Язык распознавания	
Язык распознавания для якорей	
Прямое извлечение текста из PDF	No
ОСR размер ячейки по горизонтали	40
ОСR размер ячейки по вертикали	20
Фильтр горизонтальных линий	5
Автоповорот страницы	True
Автовыравнивание страницы	True
Критерий длины линии. Автовыравнив	ar True
Обрабатывать Pdf аннотации	False
Обрабатывать все страницы	

При автовыравнивании страницы распознается самая длинная ОСК линия, имеющаяся на странице (чаще всего это линия из таблицы или атрибута) и определяет угол ее наклона относительно горизонтали. После этого происходит разворот документа, чтобы найденная линия стала параллельна линии горизонта.

По умолчанию значение параметра «True», рекомендуем оставлять его без изменения.

Этот параметр нежелательно использовать в случае, если сканированный документ сильно повернут и угол поворота составляет более 40, в этом случае Шаблонизатор не сможет распознать, в какую сторону нужно выровнять документ и выдаст ошибку. В такой ситуации можно использовать параметр «Корректировка угла».

Критерий длины. Автовыравнивание страницы

Критерий длины. Автовыравнивание страницы – параметр, позволяющий ограничивать длину линии. Данный параметр используется как делитель для ширины страницы. Если длина линии меньше вычисленного значения, то она



пропускается в алгоритме. Если страница изначально сильно повернута, то следует указать или большее значение или 0 для отключения параметра. По умолчанию значение равно 10.

łастройки	
OCR модуль	TesseractOCR
ОСЯ масштаб	2
Язык распознавания	
Язык распознавания для якорей	
Прямое извлечение текста из PDF	No
ОСR размер ячейки по горизонтали	40
ОСВ размер ячейки по вертикали	20
Фильтр горизонтальных линий	5
Автоповорот страницы	True
Автовыравнивание страницы	True
Критерий длины линии. Автовыравнивание страницы	10

Обрабатывать PDF аннотации

Обрабатывать PDF аннотации – параметр, позволяющий включить обработку аннотаций PDF-файла.

OCR модуль	TesseractOCR
OCR масштаб	2
Язык распознавания	
Язык распознавания для якорей	
Прямое извлечение текста из PDF	No
ОСЯ размер ячейки по горизонтали	40
ОСR размер ячейки по вертикали	20
Фильтр горизонтальных линий	5
Автоповорот страницы	True
Автовыравнивание страницы	True
Критерий длины линии. Автовыравнивание	10
Обрабатывать Pdf аннотации	False 🗸
Обрабатывать все страницы	True
Склеивать блоки	False
Разлелить блоки	

Обрабатывать все страницы

Обрабатывать все страницы – параметр, позволяющий включить обработку всех страниц в документе. Этот режим не делает проверку якоря LastPage.



lастройки	
OCR модуль	TesseractOCR
OCR масштаб	2
Язык распознавания	
Язык распознавания для якорей	
Прямое извлечение текста из PDF	No
OCR размер ячейки по горизонтали	40
OCR размер ячейки по вертикали	20
Фильтр горизонтальных линий	5
Автоповорот страницы	True
Автовыравнивание страницы	True
Критерий длины линии. Автовыравнивание	10
Обрабатывать Pdf аннотации	False
Обрабатывать все страницы	False 🗸
Склеивать блоки	True
Разделить блоки	False
Иснорировать онибии акорей	

Склеивать блоки

Склеивать блоки – параметр, позволяющий склеивать рядом стоящие блоки в один блок.

ОСК масштаб 2 Язык распознавания Язык распознавания для якорей Прямое извлечение текста из PDF No ОСК размер ячейки по горизонтали 40 ОСК размер ячейки по вертикали 20 Фильтр горизонтальных линий 5 Автоповорот страницы Тгце Автовыравнивание страницы True Критерий длины линии. Автовыравнивание 10 Обрабатывать Рdf аннотации False Окневать блоки False Разделить блоки True	OCR модуль	TesseractOCR
Язык распознавания Язык распознавания для якорей Прямое извлечение текста из PDF No ОСК размер ячейки по горизонтали 40 ОСК размер ячейки по вертикали 20 Фильтр горизонтальных линий 5 Автоповорот страницы True Автовыравнивание страницы True Критерий длины линии. Автовыравнивание 10 Обрабатывать Pdf аннотации False Обрабатывать все страницы False Склеивать блоки True	OCR масштаб	2
Язык распознавания для якорей Прямое извлечение текста из PDF No ОСR размер ячейки по горизонтали 40 ОСR размер ячейки по вертикали 20 Фильтр горизонтальных линий 5 Автоповорот страницы True Автовыравнивание страницы True Критерий длины линии. Автовыравнивание 10 Обрабатывать Pdf аннотации False Обрабатывать все страницы False Склеивать блоки True	Язык распознавания	
Прямое извлечение текста из PDF No OCR размер ячейки по горизонтали 40 OCR размер ячейки по вертикали 20 Фильтр горизонтальных линий 5 Автоповорот страницы True Автовыравнивание страницы True Критерий длины линии. Автовыравнивание 10 Обрабатывать Pdf аннотации False Обрабатывать все страницы False Склеивать блоки False Разделить блоки True	Язык распознавания для якорей	
ОСR размер ячейки по горизонтали 40 ОСR размер ячейки по вертикали 20 Фильтр горизонтальных линий 5 Автоповорот страницы True Автовыравнивание страницы True Критерий длины линии. Автовыравнивание 10 Обрабатывать Pdf аннотации False Обрабатывать все страницы False Склеивать блоки False Разделить блоки True	Прямое извлечение текста из PDF	No
ОСR размер ячейки по вертикали 20 Фильтр горизонтальных линий 5 Автоповорот страницы Тгце Автовыравнивание страницы Тгце Критерий длины линии. Автовыравнивание 10 Обрабатывать Pdf аннотации False Обрабатывать все страницы False Склеивать блоки False	ОСR размер ячейки по горизонтали	40
Фильтр горизонтальных линий 5 Автоповорот страницы Тгце Автовыравнивание страницы Тгце Критерий длины линии. Автовыравнивание 10 Обрабатывать Pdf аннотации False Обрабатывать все страницы False Склеивать блоки False	ОСR размер ячейки по вертикали	20
Автоповорот страницы Тгче Автовыравнивание страницы Тгче Критерий длины линии. Автовыравнивание 10 Обрабатывать Pdf аннотации False Обрабатывать все страницы False Склеивать блоки False Разделить блоки Тгче	Фильтр горизонтальных линий	5
Автовыравнивание страницы Тгче Критерий длины линии. Автовыравнивание 10 Обрабатывать Pdf аннотации False Обрабатывать все страницы False Склеивать блоки False Разделить блоки Тгче	Автоповорот страницы	True
Критерий длины линии. Автовыравнивание 10 Обрабатывать Pdf аннотации False Обрабатывать все страницы False Склеивать блоки False Разделить блоки Тгце	Автовыравнивание страницы	True
Обрабатывать Pdf аннотации False Обрабатывать все страницы False Склеивать блоки False Разделить блоки Тгце	Критерий длины линии. Автовыравнивание	10
Обрабатывать все страницы False Склеивать блоки False Разделить блоки Тгие	Обрабатывать Pdf аннотации	False
Склеивать блоки False Разделить блоки Тгие	Обрабатывать все страницы	False
Разделить блоки Тгие	Склеивать блоки	False
	Разделить блоки	True

Разделить блоки

Разделить блоки – параметр, позволяющий разделить блоки содержащие пробелы на блоки без пробелов.

🚯 Sherpa RPA

стройки	
OCR модуль	TesseractOCR
OCR масштаб	2
Язык распознавания	
Язык распознавания для якорей	
Прямое извлечение текста из PDF	No
OCR размер ячейки по горизонтали	40
ОСR размер ячейки по вертикали	20
Фильтр горизонтальных линий	5
Автоповорот страницы	True
Автовыравнивание страницы	True
Критерий длины линии. Автовыравнивание	e 10
Обрабатывать Pdf аннотации	False
Обрабатывать все страницы	False
Склеивать блоки	False
Разделить блоки	False
Игнорировать ошибки якорей	True
Удалить блоки превышающие размер	False
Процент изображения для распознавания	

Игнорировать ошибки якорей

Игнорировать ошибки якорей – параметр, позволяющий отключить генерацию ошибок, если якори (любые) не найдены. В этом случае, область якоря считается нулевой.

OCR модуль	TesseractOCR
OCR масштаб	2
Язык распознавания	
Язык распознавания для якорей	
Прямое извлечение текста из PDF	No
ОСR размер ячейки по горизонтали	40
ОСR размер ячейки по вертикали	20
Фильтр горизонтальных линий	5
Автоповорот страницы	True
Автовыравнивание страницы	True
Критерий длины линии. Автовыравнивание	10
Обрабатывать Pdf аннотации	False
Обрабатывать все страницы	False
Склеивать блоки	False
Разделить блоки	False
Игнорировать ошибки якорей	False 🗸
Удалить блоки превышающие размер	True
Процент изображения для распознавания	False
Найти печать	Faise

Удалить блоки превышающие размер

Удалить блоки превышающие размер – параметр, позволяющий удалить блоки превышающие указанный размер. Поле для ввода расположено рядом с наименованием параметра. Если указать одно число, то будут удалены блоки, у которых ширина или высота будут больше этого значения. Также, можно указать значения через запятую в формате: ширина, высота.



стройки	
OCR модуль	TesseractOCR
OCR масштаб	2
Язык распознавания	
Язык распознавания для якорей	
Прямое извлечение текста из PDF	No
ОСЯ размер ячейки по горизонтали	40
ОСЯ размер ячейки по вертикали	20
Фильтр горизонтальных линий	5
Автоповорот страницы	True
Автовыравнивание страницы	True
Критерий длины линии. Автовыравнивание	10
Обрабатывать Pdf аннотации	False
Обрабатывать все страницы	False
Склеивать блоки	False
Разделить блоки	False
Игнорировать ошибки якорей	False
Удалить блоки превышающие размер	

Процент изображения для распознавания

Процент изображения для распознавания – параметр, позволяющий указать процент изображения, который пойдет на распознавание ОСR движком. Поле для ввода расположено рядом с наименованием параметра. Значение необходимо записывать как одно число или два числа через тире.

Например:

30 (0-30) - будет распознано 0-30% изображения;

30-70 - будет распознана часть изображения;

70-100 - будет распознано нижние 30% изображения.

OCR модуль	TesseractOCR
OCR масштаб	2
Язык распознавания	
Язык распознавания для якорей	
Прямое извлечение текста из PDF	No
ОСR размер ячейки по горизонтали	40
ОСR размер ячейки по вертикали	20
Фильтр горизонтальных линий	5
Автоповорот страницы	True
Автовыравнивание страницы	True
Критерий длины линии. Автовыравнивание	10
Обрабатывать Pdf аннотации	False
Обрабатывать все страницы	False
Склеивать блоки	False
Разделить блоки	False
Игнорировать ошибки якорей	False

Найти печать



Найти печать – параметр, позволяющий включить поиске печатей на документе. В атрибуты будет добавлен ключ "Stamps", который возвратит массив объектов StampItem со свойствами X, Y, Width, Height, PageIndex.

ОСК масштаб 2 Язык распознавания	OCR модуль	TesseractOCR	
Язык распознавания Язык распознавания для якорей Прямое извлечение текста из PDF No OCR размер ячейки по горизонтали 40 OCR размер ячейки по горизонтали 20 Филь тр горизонтальных линий 5 Автоповорот страницы True Автовыравнивание страницы True Критерий длины линии. Автовыравнивание 10 Обрабатывать все страницы False Склеивать блоки False Разделить блоки False Удалить блоки превышающие размер False Преморажения для распознавания False	OCR масштаб	2	
Язык распознавания для якорей Прямое извлечение текста из PDF No ОСR размер ячейки по горизонтали 40 ОСR размер ячейки по вертикали 20 Фильтр горизонтальных линий 5 Автоповорот страницы True Автовыравнивание страницы True Обрабатывать Рdf аннотации False Обрабатывать Pdf аннотации False Склеивать блоки False Разделить блоки False Удалить блоки превышающие размер False Процент изображения для распознавания False	Язык распознавания		
Прямое извлечение текста из PDF No ОСR размер ячейки по горизонтали 40 ОСR размер ячейки по вертикали 20 Филь тр горизонтальных линий 5 Автоповорот страницы True Автовыравнивание страницы True Обрабатывать Рdf аннотации False Обрабатывать все страницы False Обрабатывать все страницы False Оклеивать блоки False Удалить блоки превышающие размер False Процент изображения для распознавания False	Язык распознавания для якорей		
ОСR размер ячейки по горизонтали 40 ОСR размер ячейки по вертикали 20 Фильтр горизонтальных линий 5 Автоповорот страницы True Автовыравнивание страницы True Критерий длины линии. Автовыравнивание 10 Обрабатывать Рdf аннотации False Обрабатывать все страницы False Склеивать блоки False Разделить блоки False Удалить блоки превышающие размер False Процент изображения для распознавания False	Прямое извлечение текста из PDF	No	
ОСR размер ячейки по вертикали 20 Фильтр горизонтальных линий 5 Автоповорот страницы True Автовыравнивание страницы True Критерий длины линии. Автовыравнивание 10 Обрабатывать Pdf аннотации False Оклеивать блоки False Разделить блоки False Удалить блоки превышающие размер False Процент изображения для распознавания False	ОСR размер ячейки по горизонтали	40	
Фильтр горизонтальных линий 5 Автоповорот страницы Тгче Автовыравнивание страницы Тгче Критерий длины линии. Автовыравнивание 10 Обрабатывать все страницы False Склеивать блоки False Разделить блоки False Удалить блоки превышающие размер False Процент изображения для распознавания False	ОСR размер ячейки по вертикали	20	
Автоповорот страницы Тгue Автовыравнивание страницы Тгue Критерий длины линии. Автовыравнивание Обрабатывать Рdf аннотации False Обрабатывать все страницы False Склеивать блоки False Разделить блоки False Игнорировать ошибки якорей False Удалить блоки превышающие размер Процент изображения для распознавания	Фильтр горизонтальных линий	5	
Автовыравнивание страницы Тrue Критерий длины линии. Автовыравнивание Обрабатывать Pdf аннотации False Обрабатывать все страницы False Склеивать блоки False Разделить блоки False Игнорировать ошибки якорей False Удалить блоки превышающие размер Процент изображения для распознавания	Автоповорот страницы	True	
Критерий длины линии. Автовыравнивание 10 Обрабатывать Pdf аннотации False Обрабатывать все страницы False Склеивать блоки False Разделить блоки якорей False Игнорировать ошибки якорей False Удалить блоки превышающие размер Процент изображения для распознавания	Автовыравнивание страницы	True	
Обрабатывать Рdf аннотации False Обрабатывать все страницы False Склеивать блоки False Разделить блоки False Игнорировать ошибки якорей False Удалить блоки превышающие размер Процент изображения для распознавания	Критерий длины линии. Автовыравниван	ие 10	
Обрабатывать все страницы False Склеивать блоки False Разделить блоки False Игнорировать ошибки якорей False Удалить блоки превышающие размер Процент изображения для распознавания	Обрабатывать Pdf аннотации	False	
Склеивать блоки False Разделить блоки False Игнорировать ошибки якорей False Удалить блоки превышающие размер Процент изображения для распознавания	Обрабатывать все страницы	False	
Разделить блоки False Игнорировать ошибки якорей False Удалить блоки превышающие размер Процент изображения для распознавания	Склеивать блоки	False	
Игнорировать ошибки якорей False Удалить блоки превышающие размер Процент изображения для распознавания	Разделить блоки	False	
Удалить блоки превышающие размер Процент изображения для распознавания	Игнорировать ошибки якорей	False	
Процент изображения для распознавания	Удалить блоки превышающие размер		
United and a second sec	Процент изображения для распознавани	ян	
Hautu nevats False	Найти печать	False	
True		True	

1.3.2 Якоря

Якоря — блоки на обр аботанном документе, с помощью которых робот определяет, в какой части документа размещаются определенные данные.

При помощи якорей задаются границы определенных атрибутов. Якорей может быть неограниченное количество, в зависимости от типа документа.

Инструменты для работы с якорями

При создании нового якоря на рабочей панели открывается список свойств и инструментов для работы с якорем.



	- 🛧 🕁 😳 🛅 Привязка: 🃚	秦 ↑ ↓
8	Якорь	1
	ID	1
	Якорные привязки	Привязки (0)
	Тип	Text
	Значение	
	Режим работы	Default
	Режим проверки значений	Default
	Раздельные слова	True
	Игнорировать регистр символов	True
	Количество блоков	0
	Проверка позиции блоков	True
	Реверсивный поиск	False
	Захватить близлежащие блоки	
	Процент изображения	
	Описание	
	Подчиненные якоря	Якоря (0)

Основные инструменты для работы с якорями.

№ п/п	Элемент интерфейса	Описание элемента интерфейса
1.	кнопка	Позволяет добавить новый якорь. Кнопка • , которая находится рядом позволяет добавить новый подчиненный якорь, при этом нужно сделать клик на свойство "Подчиненные якоря" у того якоря, которому нужно добавить подчиненный якорь
2.	кнопка 🛧	Позволяет переместить якорь вверх.
3.	кнопка	Позволяет переместить якорь вниз.
4.	кнопка	Позволяет копировать выделенный якорь и автоматически добавить новый якорь вместе с заданными настройками.
5.	кнопка 🛅	Позволяет удалить выделенный якорь.
6.	кнопка 🕏	Позволяет добавить новую привязку к выделенному якорю.

🚯 Sherpa RPA

7.	кнопка 🐲	Позволяет удалить выделенную привязку.
8.	кнопка	Позволяет переместить привязку вверх.
9.	кнопка	Позволяет переместить привязку вниз.



Настройки якоря

Параметры настроек якоря

№ п/п	Параметр	Значение параметра
1.	ID	Идентификатор якоря. Присваивается
		автоматически.
2.	Якорные	Список якорных привязок.
	привязки	
3.	Тип	Тип якоря. Выбирается из выпадающего списка:
		 Text - текст; Table - таблица или ячейка таблицы; Packet - пакет взаимозаменяемых якорей (при привязке к этому якорю будут взяты его подчиненные якоря, первый успешно найденный из списка).

Sherpa RPA

4.	Значение	Значение якоря. Для якоря типа Text возможно указывать подстановочные знаки (*)(?) Для якоря типа Table значение записывается в формате «Индекс таблицы:Индекс строки:Индекс колонки». Например, для Text «*обл* страх*» (два слова при включении «Раздельные слова»), для Table – «0:1:2» (таблица 0, строка 1, колонка 2), «1:2» (таблица 1, строка 2), «2» (вся таблица 2)
5.	Режим работы	 Режим работы якоря. Возможные значения: Default – якорь обязателен на первой странице многостраничного документа; LastPage – якорь будет обязателен в многостраничном документе и укажет на последнюю страницу; Optional – необязательный якорь и если он не будет найден на странице, то ошибка не генерируется.
6.	Режим проверки значения	 Выбор режима проверки значений. Возможные значения: Default (по умолчанию); EveryWordContains (проверка каждого слова в значении как вхождение в строку. Пример: "тест" => "*тест*"); RemoveNonWordSumbols (при проверке удаляются все не числовые и буквенные символы).


7.	Раздельные слова	При включенной настройке значение якоря разбивается на слова и каждое слово ищется отдельно.
8.	Игнорировать регистр символов	При включенной настройке будет игнорироваться регистр символов.
9.	Количество блоков	Если включена настройка «Раздельные слова», то при поиске якоря возможна ситуация, когда в заданной области могут быть найдены несколько значений блоков, соответствующие заданным словам. Данное свойство позволяет ограничить максимальное количество слов между блоками.
10.	Проверка позиции блоков	При включенной настройке проверяется позиция блоков относительно координаты Y и берутся блоки с наименьшим значением.
11.	Реверсивный поиск	При включенной настройке поиск якоря будет идти снизу страницы.
12.	Захватить близлежащие блоки	Захват ближних блоков якоря. Позволяет захватить в якорь ближние блоки от уже захваченного значения. Например, если указать +2-1, то к уже захваченному якорю будут присоединены ближние блоки: 2 блока справа и 1 блок слева. Данное свойство может использоваться, когда не удается установить полное устойчивое значение для якоря, но часть его имеет устойчивое значение.

13.	Процент изображения	Процент изображения, в котором будет производиться поиск якоря. Значение необходимо записывать как одно число или два числа через тире. Например: 30 (0-30) - будут использоваться 0-30% изображения; 30-70 - будут использоваться часть изображения; 70-100 - будут использоваться нижние 30%
		изображения.
14.	Описание	Пользовательское описание якоря.
15.	Подчиненные якоря	Якоря, напрямую зависящие от родительского якоря. Поиск подчиненных якорей будет происходит только после того, как будет найден главный/родительский якорь. В случае, если в настройках родительского якоря указан режим «Optional» (т.е. он не является обязательным), и если родительский якорь не был найден, то и поиск подчиненных якорей не будет осуществлен, они будут пропускаться. Эта параметр используется, если атрибут находится в середине многостраничного документа.

1.3.3 Атрибуты

Атрибуты позволяют определить и обозначить области документа, данные из которых будут распознаваться роботом и импортироваться в таблицу «Результаты».

Основные инструменты для работы с атрибутами.

№ п/п	Элемент интерфейса	Описание элемента интерфейса
1.	кнопка 🕂	Позволяет добавить новый атрибут.
2.	кнопка 🔨	Позволяет переместить атрибут вверх.
3.	кнопка♥	Позволяет переместить атрибут вниз.
4.	кнопка	Позволяет копировать выделенный атрибут.
5.	кнопка	Позволяет удалить выделенный атрибут.
6.	кнопка 🕏	Позволяет добавить новую привязку к выделенному атрибуту.
7.	кнопка 📚	Позволяет удалить выделенную привязку.
8.	кнопка 🔨	Позволяет переместить привязку вверх.
9.	кнопка	Позволяет переместить привязку вниз.



Свойства атрибутов

При создании нового атрибута на рабочей панели открывается список свойств и инструментов для работы с атрибутом.

Свойства и инструменты для работы с атрибутом.

№ п/п Элемент интерфейса	Описание элемента интерфейса
--------------------------	------------------------------

1.	Имя	Имя атрибута, которое будет отражено в результатах
2.	Якорные привязки	Список якорных привязок атрибута.
3.	Ограничение линиями	При включенной настройке осуществляется поиск ограничения линиями. Атрибут может быть ограничен с любой стороны, относительно линий, распознанных ОСК модулем. Эти линии отображаются при нажатии кнопки «ОСК Линии». Возможные значения: • None – отключено; • UpDown – сверху и снизу; • Up – сверху; • Down – снизу.
4.	Корректировать угол	При включенной настройке осуществляется корректировка угла наклона области. Область корректируется, если она соприкасается с границами таблицы. При включенном значении «Автовыравнивание» корректировку угла использовать не нужно.
5.	Привязка области	Привязка к области осуществляется по 4-м параметрам: координата X, координата Y, ширина, высота области. Исходная точка области - вверху слева.

6.	Область	Прямоугольник области атрибута на странице.
7.	Мин количество символов	 Минимальное количество символов, содержащихся в атрибуте. По умолчанию указано значение 0. Настройка используется, если атрибут должен содержать в себе определенное количество символов, например, атрибут ИНН. Если вычисленное значение меньше указанного будет сгенерирована ошибка.
8.	Тип данных	 Тип данных для атрибута. Возможные значения: Техt - текст; Int - число; Float - вещественное число; Date Time - дата и время; Chars - только буквы (исключаются цифры и другие символы); CountColorHSV - количество точек в заданном интервале цвета. Значение должно быть указано в формате «H1-H2;S1-S2;V1-V2», где H1, S1, V1 - это минимальные значения, а H2, S2, V2 - максимальные значения; Regex - регулярное выражение.

9.	Формат значения	Формат значения атрибута. Задается форматирование значения. Например: • если тип значения дата и время то, чтобы получить только дату, необходимо указать «dd.MM.yyyy»; • если тип значения CountColorHSV, то необходимо указать интервал цвета в формате «H1-H2;S1-S2;V1- V2». Формат указывается без кавычек; • Если тип значения Regex, то необходимо указать регулярное выражение.
10.	Ошибка если значение пустое	Если при обработке шаблона значение атрибута пустое, то будет сгенерирована ошибка распознавания.
11.	Проверять якорные привязки	Проверка якорных привязок атрибута. При включенной настройке будет производиться проверка якорных привязок. Если хоть одна из привязок не будет найдена, то атрибут, также, не будет найден.
12.	Режим работы	Режим работы атрибута. Возможные значения: • Default - стандартная работа атрибута (его значение будет заполнено при первом успешном определении якорей);



		 Аррепd - добавление значения в атрибут при каждом успешном определении якорей на следующих страницах; Overwrite - перезапись, а не добавление значения атрибута.
13.	Описание	Пользовательское описание атрибута.

Ւ 🕁 📴 📋 Привязка: 🤹	
рибут	Без имени
Имя	Без имени
Якорные привязки	Привязки (0)
Ограничение линиями	None
Корректировать угол	False
Привязка области	
Область	0:0:0:0
Мин количество символов	0
Тип данных	Text
Формат значения	
Ошибка если значение пустое	False
Проверять якорные привязки	False

1.3.4 Таблицы

На рабочей панели справа отображаются все таблицы шаблона, которые были распознаны ранее. Их может быть неограниченное количество. Каждой таблице автоматически присваивается номер, нумерация начинается с 0. Также, отображается количество колонок каждой распознанной таблицы. Именно по этому значению можно определить с какой (или с какими) таблицами продолжить работу.

Основные инструменты для работы с таблицами.

№ п/п	Элемент интерфейса	Описание элемента интерфейса
1.	кнопка +	Позволяет добавить новую таблицу.

2.	кнопка	Позволяет копировать выделенную таблицу.
3.	кнопка	Позволяет удалить выделенную таблицу.
4.	кнопка 🕏	Позволяет привязать якорь к выбранному свойству.
5.	кнопка 📚	Позволяет удалить якорь из выбранного свойства.
6.	кнопка +	Позволяет добавить колонку таблицы.
7.	кнопка 🛅	Позволяет удалить колонку таблицы.
8.	кнопка Đ	Импортировать все таблицы из документа.



сходный фай	n 🕨 2) 06p	аботать документ — 🕨 Пров	ерипь (+3) Добавить а	трибут 🍙 4) Вы	полнить п	олное сканирование 🔞 Умный г	помощник		
• Масшл	a6: 35%	- @ @ Orofpace	аты: Эякора ЭАтриб	The Talinana	@ Hepres	ки таблиц 🕲 ОСК блоки 🕲 ОСК	Tationalas @ OCR./	umutur	Табляцы
	Мы, нися И	еподписаециеся, Заместите Публичное анц	взаими и Публя ль генерального дире рюнерное общество "З	Акт с ых расчетов за п между АО теное акционерно по договору 2023 стора по аконов јемланика"	верки мериод 9 г Матана ре общест 802 от 03. вике и фи	месяцев 2023 г. 1° 1° Земляника" 02 2022 изансам АО "Матена" Иванов с другой сторою	Иван Иванович. 4. составиля насто	одной стороны, нанй акт сверня	+ Q) () S S Koover + ()
	B TOM, 47	о состояние взаимных расчет	гов по данным учета сл	едующее:		- 1954 AUMOR/2018497			
	По данным	AO 'Manuka', pyő.			По данны	м публичного акционерного общест	тва "Эемпяника"		
	Data	Документ	Дебет	Kpegart	(Jara	Документ	Deter	Konpert	
	Campao MP	42014404 December (57 or 03 02 2023)	108 193 92		Сальдо н	VETWOR			
	21.02.23	Onnata (865 of 21.02.2023)	100 1000	108 193.93					
	28.02.23	Продежа (119 от 28.02.2023)	108 193,92						
	31.03.23	Ponexia (176 or 31.03.2023)	108 193 92	108 193.92					
	25.04.23	Onnata (2412 of 25.04.2023)	100.770.78	108 193,92					
	30.04.23	Продяжа (242 от 30.04.2023)	108 193,92						
	31.65.23	Dissessing (306 or 25.05.2023)	106 193 62	108 193.92					
	27.06.23	Omnava (3664 or 27.06.2023)	121.12574	108 193,92					
	30.06.23	Продежа (369 от 30.06.2023)	108 193,92						
	31.07.23	Teodesia (425 or 31.07.2023)	108 195.92	100 190,92					
	29 58 23	Cinnana (4969 or 29.06 2023)		108 193.90					
	51.08.2.5	Ppogawa (500 or 31.06.2023)	108 193,92						
	30.09.23	Doossea (561 or 30.09.2023)	108 195 92	100,190,94					
	Обороты	a nepeog	973 745,28	865 551,36	Обороты	а период			
	Сальдо ко	ne-suoe	108 193,92		Сальдо ко	scarga			
					On partners	Definence assessment of user tao "3	escoleman"		
	По данным	AO 'Matera"	0.070/023						
	108 103 92	23 SADORREPHOCTS & HORBAY AO "M	tanena" orto teus nofins 92 someñas)						
	OT AO 'NA	tatwig"			Ot Tlydnese	ное акционерное общество "Земплиия	¢.		Tafinesa
	-		the second second second						Таблица 0 (Козонка 8)
	Same Inter	с тенерального деректора по эконом	and is destinent		-				
	22		(Ensure K.K.)				C		
	-		(cannot strong						
	and the second s								
									the standard
									👍 Beops
									Т Ацибиъ
									III Tatavisi



Свойства таблиц

На рабочей панели открывается список свойств и инструментов для работы с таблицей.

Свойства и инструменты для работы с таблицами.

№ п/п	Элемент интерфейса	Описание элемента интерфейса
1.	Индекс	Индекс ОСК таблицы. Для улучшения привязки рекомендуется указывать минимальное количество колонок. В
		случае, если Шаблонизатор извлекает данные лишь из одной таблицы сканированного документа, необходимо указывать индекс, равный 0.
2.	Пропустить строки	При необходимости возможно указать количество первых строк, которые нужно пропустить для того, чтобы данные этих строк не попали в результаты работы робота. При выборе значения, равного 1, — первая строка будет пропущена. Наименования колонок будут соответствовать значениям первой строки. Если задать названия колонок в шаблоне, то эти значения не будут изменены.
3.	Строк в заголовке	При построении таблицы заданное количество строк будет использовано для создания заголовков колонок. При указании значения 0, названия будут взяты из шаблона колонок, в противном случае название будет суммироваться из ячеек

		колонки по заданному количеству строк. Например, если в заголовок должна попасть информация, размещенная в двух верхних строках таблицы, — необходимо указать количество строк в заголовке, равное 2.
4.	Связующий якорь	При работе с документом, в котором есть несколько таблиц можно указать связующий якорь, имеющий тип Optional. Якорь укажет на какие-либо слово, заголовок, пункт, после которых начинается нужная таблица. При нахождении такого якоря таблица, следующая за ним, будет распознаваться.
5.	Финишный якорь	Финишный якорь ограничивает извлечение таблицы до якоря, который может иметь тип Optional. Таким якорем, например, может служить строка в таблице "Всего к оплате". Таблица будет извлечена из документа до финишного якоря.
6.	Верхний ограничивающий якорь	Якорь, после которого таблица начнет строиться.
7.	Нижний ограничивающий якорь	Якорь, до которого таблица будет строиться.
8.	Минимальное количество колонок	Минимальное количество колонок в ОСК таблице. Все таблицы с количеством

		колонок меньше указанного пропускаются.
9.	Максимальное количество колонок	Максимальное количество колонок в ОСК таблице. Все таблицы с количеством больше указанного пропускаются.
10.	Точное равенство колонок	При включенной настройке извлекаться будут только те таблицы, у которых количество колонок в документе меньше или равно заданному в шаблоне.
11.	Колонки на первой странице	При включенной настройке колонки будут браться только с первой страницы. Данная настройка указывает на то, что на следующих страницах у таблицы нет заголовков.
12.	Склеивать таблицы	При включенной настройке таблицы с одинаковым количеством колонок будут склеиваться. Если настройка не включена, такие таблицы будут добавляться как новые.
13.	Ошибка если таблица пустая	При обработке шаблона, если таблица пустая, то будет сгенерирована ошибка распознавания.
14.	Колонки	Колонки таблицы. По каждой таблице отображается соответствующее ей количество колонок. Также, здесь будут указываться: • Индекс (задается автоматически);

• Имя (задается пользователем);
• Тип данных (Text – текст; Int –
число; Float – вещественное число; Date
Time – дата и время; Chars – только буквы
(исключаются цифры и другие символы)).
В случае, если в результатах полного
сканирования отсутствует несколько
колонок, то мы можем их добавить.

Индекс О Пропустить строки 1 Строк в заголовке 1 Связующий якорь 1 Финишный якорь 1 Верхний ограничивающий якорь 1 Нижний ограничивающий якорь 1 Минимальное количество колонок 0 Точное равенство колонок 6 Колонки на первой странице False Склеивать таблицы True Ошибка если таблица пустая False	-	Табрица	
Пропустить строки 1 Строк в заголовке 1 Связующий якорь 1 Финишный якорь 1 Верхний ограничивающий якорь 1 Нижний ограничивающий якорь 1 Минимальное количество колонок 0 Точное равенство колонок False Колонки на первой странице False Склеивать таблицы True Ощибка если таблица пустая False	Ţ	Индекс	0
Строк в заголовке 1 Связующий якорь Финишный якорь Верхний ограничивающий якорь Нижний ограничивающий якорь Минимальное количество колонок 0 Максимальное количество колонок 0 Точное равенство колонок False Колонки на первой странице False Склеивать таблицы Тrue Ощибка если таблица пустая False		Пропустить строки	1
Связующий якорь Финишный якорь Верхний ограничивающий якорь Нижний ограничивающий якорь Минимальное количество колонок 0 Максимальное количество колоно 0 Точное равенство колонок False Колонки на первой странице False Склеивать таблицы Тrue Ощибка если таблица пустая False		Строк в заголовке	1
Финишный якорь Верхний ограничивающий якорь Нижний ограничивающий якорь Минимальное количество колонок 0 Максимальное количество колоно 0 Точное равенство колонок False Колонки на первой странице False Склеивать таблицы Тrue Ощибка если таблица пустая False		Связующий якорь	
Верхний ограничивающий якорь Нижний ограничивающий якорь Минимальное количество колонок 0 Максимальное количество колоно 0 Точное равенство колонок False Колонки на первой странице False Склеивать таблицы Тrue Ощибка если таблица пустая False		Финишный якорь	
Нижний ограничивающий якорь Минимальное количество колонок 0 Максимальное количество колоно 0 Точное равенство колонок False Колонки на первой странице False Склеивать таблицы Тrue Ощибка если таблица пустая False		Верхний ограничивающий якорь	
Минимальное количество колонок 0 Максимальное количество колонос 0 Точное равенство колонок False Колонки на первой странице False Склеивать таблицы True Ощибка если таблица пустая False		Нижний ограничивающий якорь	
Максимальное количество колонс 0 Точное равенство колонок False Колонки на первой странице False Склеивать таблицы True Ошибка если таблица пустая False		Минимальное количество колонок	0
Точное равенство колонок False Колонки на первой странице False Склеивать таблицы True Ошибка если таблица пустая False		Максимальное количество колоно	0
Колонки на первой странице False Склеивать таблицы Тгие Ошибка если таблица пустая False		Точное равенство колонок	False
Склеивать таблицы Тгче Ошибка если таблица пустая False		Колонки на первой странице	False
Ошибка если таблица пустая False		Склеивать таблицы	True
		Ошибка если таблица пустая	False
Колонки (0)		Колонки	Колонки (0)

1.3.5 Чертежи таблиц

Чертежи таблиц применяются для документов, в которых по каким-либо причинам отсутствуют границы таблиц. Это могут быть иностранные документы (например, Invoice), в которых линии таблиц не предусмотрены. Также это могут быть российские документы, в которых частично не прорисованы линии таблицы, как на скриншоте ниже (отсутствуют межстрочные линии).



ить исходный фа	йл 🕨 2) Обработать документ — 🕨 Прове	рить (+3) Добавить а	грибут 🔁 4) Вы	полнить п	юлное сканири	ование 🛛 🛞 Умний по	мощник			
Macu	паб: 35% - 🖉 🔎 Отобража	вы: 💿 Якоря 💿 Атрибу	ты 🐼 Таблицы	•Черте	жи таблиц 🖗	OCR блоки 🗞 OCR т	абланды 🕸 ОСК Л	инии	Чертежи заблиц	
									🕂 🛉 Колонка: 📚 🏂 Ширина к	олонки 🖏
		взаимне и Публи г	Акт с ых расчетов за п между АС чное акционерн ю договору 202	сверки териод 9 О "Мапин юе общес 302 от 03	месяцев 202: а° тво "Земпяны 02.2022	аr. жа"			Veprex refinatel Onscalve Pessit represel yron Tpassit tepresel yron Tpassit received yron Tpassit received yron Tpassit received yron Tpessit received yron Pepseve opparements taffatage	Tests senses 1 TopLeft > None 2 TopFlight > None 0 BottomRight > None 0 BottomLeft > None
	Мы, нижеподписавшиеся, Заместител	ь генерального дирея	тора по эконов	иже и ф	инансам АО	"Малина" Иванов И	ван Иванович, о	одной стороны,	 Наконое опранителие таблицы	
	и Публичное акци	онерное общество "З	емпяника"			_, с другой стороны,	составили насто	ящий акт сверки	 Колончи	Привежи (0)
	в том, что состояние взаимных расчето	в по данным учета сли	st0xomfee:						 ширина колонок	5.6 (21,5 (11,5 (11,4 (5.7 (21,5)
	По данные АО "Малина", руб.			По данна	ым публичного	акционерного общество	а "Земляника"		 Publipsats crock	False
	Line I Decent	Seler	Energy	1.000		TO GOOD AND TO GOO	1999	1 100000	 > Отображение рамок	True: True: True: True
	Сальдо начальное	Meloer.	Caragers .	Camago w		Marine Contraction of the Contra	Mades	00000	 Толцина рамки	2
	03.02.23 Продина (57 or 03.02.2023)	108 193,92							 Црет рамки	#000000
	21.02.23 Onviana (865 or 21.02.2023)		108 193,96	2					 Обрамление колонок	True
	28.02.23 Opogawa (119 of 28.02.2023)	108 193,92		-				-	Обранление строк	True
	16.03.23 Onneta (1421 of 16.03.2023)		108 193,96	2	-				Обранление заголовка	True
	31,03,23 Предежа (176 от 31,03,2023)	108 193,92	100 100 0	-					 Колонки для определения строк	
	25.54.23 Onnetta (2412 07.25.54.2023)	104 105 10	108 190,0	-	-				 Привяска горизонтальной личим	Center
	25.05.23 (popular (2056 or 26.05.2023)	100 100,92	108 103 0		-				 Очещение горизонтальной тинии	0
	31.05.23 Decemen (306 or 31.05.2023)	108 193 92	100 100,0	-	-				 Криторий поиска поризонтатыных р	0
	27.06.23 Onnata (3664 or 27.06.2023)		108 193.90	2					 Рестигивать табляку	False
	30.06.23 (Ipogetea (369 or 30.06.2023)	108 193,92								
	25.07.23 Onneta (4220 of 25.07.2023)		108 193,90	2						
	31.07.23 Ppogawa (435 of 31.07.2023)	108 195,92								
	29.08.23 Onnaya (4969 of 29.08.2023)		108 193,90	2	-					
	31.06.23 (1004044 (500 01 31.06.2023)	108 193,92	108 103 0	-	-			-		
	30.09.23 Chotean (SEL or 30.09.2023)	108 103 02	108 190.0	-	-			-		
	Ofoportulas neosoa	973 745,28	865 551.30	6 Officerta	a necesor					
	Сальдо венечное	108 193,92		Сальдо в	and vision					
	По данныя АО "Матина" на 30.62.923 задолженность в пользу АО "Ма 168 193,52 руб. (Сто восемь тысяч сто девено От АО "Нарына"	лина" то три рубля 92 колейна).		По данныя От Публич	и Публичкое ак	ионерное общество "Зем в общество "Земляника"	anang,		Растягивать таблицу	
	Заместитеть генерального директора по эконом	NO & GHORNEAM						1	При включения чертех таблицы будит расти	гиваться на следницие страни.
	54 (SA (SA	12 I						100 C	 Colline in according to an original sector	to ecile thread he repeat closes
		- (Reason R.R.)		-	1	(-	 	
	an .			an a				I I		
	mire.	1 1		M.S.C.	1			I I		
		1 1			1			I I		
		1 1			1			I I	 Anna and a second se	
		1 1			1			I I	 Ф Настрайки	
		1 1			1			I I	 	
		1 1			1			1	The sector	
		1 1			1			1	T Americana	
		1 1			1			1	-	
		1			1			I	 Tadivusi	
							1			

Основные инструменты для работы с чертежами таблиц.

№ п/п	Элемент интерфейса	Описание элемента интерфейса
1.	кнопка +	Позволяет добавить чертеж таблицы.
3.	кнопка 🛍	Позволяет удалить чертеж таблицы.
4.	кнопка 🕏	Позволяет добавить привязку колонки.
5.	кнопка 📚	Позволяет удалить привязку колонки.
6.	кнопка	Позволяет добавить ширину колонки.

Чертежи таблиц		
🕂 💼 Колонка: 🃚 📚	Ширина колонки:	***

На рабочей панели справа отображаются все параметры, которые позволяют произвести настройку чертежа таблицы.



•]	Ū	ј 🛛 Колонка: 🎓 🕏 🖓 Ширина	колонки: 🎉
1	Чер	отеж таблицы	Без имени
		Описание	
	>	Левый верхний угол	0 : TopLeft -> None
	>	Правый верхний угол	0 : TopRight -> None
	>	Правый нижний угол	0 : BottomRight -> None
	>	Левый нижний угол	0 : BottomLeft -> None
		Верхнее ограничение таблицы	
		Нижнее ограничение таблицы	
		Колонки	Привязки (0)
		Ширина колонок	
		Фильтровать колонки	False
		Фильтровать строки	False
	>	Отображение рамок	True;True;True;True
		Толщина рамки	2
		Цвет рамки	#000000
		Обрамление колонок	True
		Обрамление строк	True
		Обрамление заголовка	True
		Колонки для определения строк	
		Привязка горизонтальной линии	Center
		Смещение горизонтальной линии	0
		Критерий поиска горизонтальных	0
		Растягивать таблицу	False

Свойства и инструменты для работы с чертежами таблиц.

№ п/п	Элемент интерфейса	Описание элемента интерфейса
1.	Описание	Описание чертежа таблицы.
2.	Левый верхний угол	Якорная привязка левого верхнего угла.
3.	Правый верхний угол	Якорная привязка правого верхнего угла.
4.	Правый нижний угол	Якорная привязка правого нижнего угла.
5.	Левый нижний угол	Якорная привязка левого нижнего угла.
6.	Верхнее ограничение таблицы	Ограничение таблицы сверху страницы. Можно указать значения в пикселях или процент, используя символ *. Например, 200*.
7.	Нижнее ограничение таблицы	Ограничение таблицы снизу страницы. Можно указать значения в пикселях или

		процент, используя символ *. Например, 200*.
8.	Колонки	Якорные привязки для колонок.
9.	Ширина колонок	Ширина колонок. Можно указать фиксированное значение или процент от всей ширины таблицы в формате " 50* ", " 10* ". Вся ширина таблицы равна 100. Если сумма значений процентов всех указанных колонок равна или больше 100, то ширина таблицы делится пропорционально указанным значениям. Иначе, последняя колонка будет заполняться автоматически до значения 100.
10.	Фильтровать колонки	При включенной настройке для распознавания таблицы будут учитываться только те линии, которые нарисованы в чертежи таблицы.
11.	Фильтровать строки	При включенной настройке для распознавания таблицы будут учитываться только те линии, которые нарисованы в чертеже таблицы.
12.	Отображение рамок	Управление отображением рамок чертежа таблицы.
13.	Толщина рамки	Толщина рамки таблицы в пикселях.
14.	Цвет рамки	Цвет рамки таблицы в формате RGB. Например, «#000000» — черный цвет, «#FF0000» — красный цвет.

15.	Обрамление колонок	При включенной настройке отображаются линии колонок на чертеже таблицы.
16.	Обрамление строк	При включенной настройке отображаются линии строк на чертеже таблицы.
17.	Обрамление заголовка	При включенной настройке отображается первая строка на чертеже страницы.
18.	Колонки для определения строк	Индексы колонок, по которым будут определяться строки. Если значение будет пустое, то будут участвовать все колонки. Индекс начинается с 0.
19.	Привязка горизонтальной линии	Параметр привязки горизонтальной линии. Возможные значения: • Center - по центру между блоками текста; • Top - верх нижнего блока текста; • Bottom - низ верхнего блока текста.
20.	Смещение горизонтальной линии	Смещение горизонтальной линии по оси Y в пикселях. Для точного смещения следует указывать значения в формате + Например, +10-20.
21.	Критерий поиска горизонтальных линий	Смещение блоков текста для определения их пересечения. По умолчанию значение равно 10.
22.	Растягивать таблицу	При включенной настройке чертеж таблицы будет растягиваться на следующие страницы. Данный параметр



необходим если заголовки таблицы есть

только на первой странице.

lep	теж	и таблиц		
ł	İ	Колонка: 🎓 📚 Ширина	колонки: 🖏	
~	Чер	теж таблицы	Без имени	
	Описание			
	>	Левый верхний угол	0 : TopLeft -> None	
	>	Правый верхний угол	0 : TopRight -> None	
	>	Правый нижний угол	0 : BottomRight -> None	
	>	Левый нижний угол	0 : BottomLeft -> None	
	Верхнее ограничение таблицы			
	Нижнее ограничение таблицы			
		Колонки	Привязки (0)	
		Ширина колонок		
		Фильтровать колонки	False	
		Фильтровать строки	False	
	>	Отображение рамок	True;True;True;True	
		Толщина рамки	2	
		Цвет рамки	#000000	
		Обрамление колонок	True	
		Обрамление строк	True	
		Обрамление заголовка	True	
		Колонки для определения строк		
		Привязка горизонтальной линии	Center	
		Смещение горизонтальной линии	0	
		Критерий поиска горизонтальных	0	
		Растягивать таблицу	False	

Для осуществления привязки границ таблицы для каждого угла чертежа необходимо указать якорь и задать параметры расположения линий относительно выбранного якоря.



Параметры якорных привязок углов таблицы.

№ п/п	Элемент интерфейса	Описание элемента интерфейса
1.	ID якоря	ID якоря, к которому происходит
		привязка.
2.	Параметр	Параметр привязки. Возможные значения:
		• TopLeft – привязка левого верхнего
		угла якоря к выбранному объекту
		привязки;
		• TopRight – привязка правого
		верхнего угла якоря к выбранному
		объекту привязки;
		• BottomLeft – привязка левого
		нижнего угла якоря к выбранному
		объекту привязки;
		• BottomRight – привязка правого
		нижнего угла якоря к выбранному
		объекту привязки;
		• Left – привязка левой грани якоря
		(посередине) к выбранному объекту
		привязки
		• Тор – привязка верхней грани якоря
		(посередине) к выбранному объекту
		привязки;
		• Right – привязка правой грани якоря
		(посередине) к выбранному объекту
		привязки;
		• Bottom – привязка верхней грани
		якоря (посередине) к выбранному
		объекту привязки;



		 LeftArea - привязка всей левой области якоря к выбранному объекту привязки; RightArea - привязка всей правой области якоря к выбранному объекту привязки; BottomArea - привязка всей нижней области якоря к выбранному объекту привязки; ТорArea – привязка всей верхней области якоря к выбранному объекту привязки; FullArea - привязка якоря со всех сторон к выбранному объекту привязки.
3.	Объект привязки	Точка или грань области атрибута для привязки выбранного параметра.
4.	Принудительно	При включенной настройке указанный параметр якоря будет задан к выбранному объекту привязки принудительно и без проверок.
5.	Обязательно	При включенной настройке ошибка данной привязки поиск всего элемента будет считаться неудачным.
6.	Смещение Х	Смещение по оси X в пикселях. Для точного смещения следует указывать значения в формате + Например, +10-20. Данная настройка используется в случае, когда после якоря указывается какой-либо

		символ (скобка или двоеточие), который не должен попасть в область атрибута.
7.	Смещение Ү	Смещение по оси Y в пикселях. Для точного смещения следует указывать значения в формате + Например, +10-20.

ертежи таблиц		
🕂 📋 Колонка: 📚 📚 Ширин	а колонки: 🎋	
 Чертеж таблицы 	Без имени	^
Описание		
 Левый верхний угол 	0 : TopLeft -> None	
ID якоря	0	
Параметр	TopLeft	
Объект привязки	None	
Принудительно	False	
Обязательно	False	
Смещение Х	0	
Смещение Ү	0	
> Правый верхний угол	0 : Top Right -> None	
> Правый нижний угол	0 : BottomRight -> None	
> Левый нижний угол	0 : BottomLeft -> None	

Для отображения рамок чертежа таблицы необходимо задать параметры отображения каждой из рамок.

Параметры отображения рамок чертежа таблицы.

№ п/п	Элемент интерфейса	Описание элемента интерфейса
1.	Левая рамка	Отображение левой рамки чертежа таблицы.
2.	Верхняя рамка	Отображение верхней рамки чертежа таблицы.
3.	Правая рамка	Отображение правой рамки чертежа таблицы.
4.	Нижняя рамка	Отображение нижней рамки чертежа таблицы.



 Отображение рамок 	True;True;True;True
Левая рамка	True
Верхняя рамка	True
Правая рамка	True
Нижняя рамка	True

2. Действия в Sherpa Designer после создания шаблона

2.1 Значения переменных

После создания шаблонов для каждого типа документа (их может быть несколько 5, 10 и т.д.), можно пойти двумя путями:

 Настроить работу сценария проекта с поочередным распознаванием по каждому шаблону, при этом добавлять в сценарий столько блоков, сколько было создано шаблонов



2) Запустить распознавание по списку шаблонов в одном блоке.

Прежде чем использовать список, необходимо его задать в Переменных. Выбираем Тип переменной List (список), задаем имя переменной и редактируем значение переменной: добавляем в список названия созданных нами шаблонов, по которым будет производиться распознавание. Например, ТОРГ12.json, УПД.json и т.д.

1 Paralana	Otemporen Distance Status Couper vi 10 - Human guarganum Trainen	- a - x
Courte Congress Congress Courte	Carotan Carota	
	Tanan Count Radee	1
Etder Opens Seberanes	 Bacipall Copy Honos Asyptotes 1^e - Partnesses 	(C) Continue
Parcelan Art Barconenas pater	X Litterature Advenue	Распознать на ваблону Ластознает заденный документ го цайкону
Protestans T2H	20000000 20000000000000000000000000000	fear ectories
Factorians Ant company	NYT 4480 K	Нонер страница
Tacces and Taccept		Ryna a duding sanda
Pactusaris CH015;		Prove a dealory autorese" (Text:Coverce) Prove dealory codoroes area participatements (Prove textitiones interfet)
Factorianamie CaPTCHW		() New York
Костолнать по шиблону	04 Grams	+ B Has Ta Heatingson year
Chemine		1 Перенник на Редакторовать.
8.9		
CANNER	No Representations 100%	
C mm	😮 (*) Overanis Company's a galla. Conventional lyder: Dispars away (anothe	

На панели свойств указываем путь к файлу PDF. Файл может быть многостраничным. Также поддерживаются файлы формата .png и .jpg.

Далее указываем путь к файлу шаблона для распознавания. При указании списка файлов шаблонов будут последовательно использоваться каждый файл, пока распознавание не закончится успешно. Если ни один из шаблонов не подойдет для документа, то будет сгенерировано исключение с ошибкой первого шаблона из списка.

Распознает заданны	ій докуме	нт по шаблон	y
Зидео инструкция			
21 Найти			×
Основные свойства	1		-E
Путь к файлу*	hh.	pdf	-
Номер страницы	1		
Путь к файлу ша	бл \$П	еременная1	
Выход			1
Переменные			
) Переменные - Ш			
) Переменные - Ш Имя	Тип	Начально	e 3HB1



Указываем путь к файлу в формате .pdf

В строке Путь к файлу шаблона указанную созданную нами переменную со списком шаблонов.

Номер страницы – 1, тогда распознавание начнется с первой страницы документа.

При запуске роботы блока «Распознать по шаблону», первым распознается документ, указанный в списке Переменной. В случае возникновения ошибки по одному из шаблонов, робот переходит к распознаванию следующего шаблона.

При использовании списка возможно увеличение скорости обработки нескольких шаблонов, если параметры распознавания страницы у этих шаблонов одинаковые.

Если параметры распознавания различны, то время распознавания документа увеличивается.

Например, если для одного из шаблонов указан ОСК-модуль Tesseract, а для другого шаблона – YandexVision (или разный масштаб или язык), то для каждого из этих шаблонов робот будет заново распознавать документ, т.к. указаны разные параметры.

Поэтому рекомендуем при создании шаблонов для типовых стандартизированных документов оставлять настройки, заданные по умолчанию и менять их только в случае необходимости.

В случае, если параметры настроек некоторых шаблонов были изменены, то рекомендуем компоновать их по группам: сначала указать несколько шаблонов с одним видом настроек, затем несколько шаблонов с другим типом настроек.

Индекс шаблона

После распознавания документа будет определен индекс шаблона.



HOID CEORCY		_
D) (recenture		
Распознать по шаблон	y	
Распознает заданный /	юкумент по шабл	ону
Видео инструкция		
Halini 21 Halini		×
Номер страницы		
Путь к файлу шабл		
, Выход		
Атрибуты	SAttributes	
Таблицы 🗡	STables	_
	STranker Continues	

Используя блок «Условие» мы задаем траекторию дальнейшего сценария согласно полученным индексам. (например, если индекс 0 – то это УПД, если индекс 1 – Счет-фактура и т.д.)

Если для распознавания использовался один документ, а не список, то индекс по умолчанию всегда будет 0 (нулевой).

Панель свойства Свойства Распознает заданный документ по шаблону Видео инструкция Выход Атрибуты \$Attributes Таблицы \$Tables Индекс шаблона \$IndexPattern

После обработки шаблона необходимо продолжить работу с созданными атрибутами.

Атрибуты отображаются в окне «Результаты полного сканирования» Шаблонизатора (*Имя, Значение*).

Атрибуты



Режим ра Атрибуты	боты:	Роким	полного ра	спознавания	• •
Имя	Зна	зчение	6		
Номер	14				

Чтобы получить значения атрибутов мы должны обратиться к словарю и присвоить значение переменной. Используем блок «Присвоить значение переменной».

На панели свойств выбираем значение переменной, кликаем на «…» и редактируем выражение. В открывшемся окне выбираем категорию \$*Attributes*.

SAttributes Sakatrink]	Редактировать в венние	- C Casiluna
Sector And construction and many reacting and construction of the interview Exector Approximation Exector A	SAttributes['Steamene'] 3	Присакить значение переменной Блок устанализает новые значения для одлой или нескольких переменник, Новое значения переманной искиет быть как канстан-той (числов, техстой) герокой), так
Sciences Description Karengese Zoosage uproveme under scheme Defensionen Berkerp Defensionen Berkerp Defensionen Berkerp Berkerp Berkerp		24 Maine X
Scherophie Zestation generation Opportunities Additional Antiportunities Contraction <	Protectorements	, Doossue coolens
Scherophie Zestage uppervinte, under bit statution SAttributes Department 2	пастрая останка	Переменная I
Dependence General procession General procession Statutores Dependence	Калегория Дзажды целкниге, чтобы вставиты	3-2-envel (1)-
Bindess Bindess Tependemine Bindesset	Orginanten er Deratgi Titteren begans titt	3 SAttributes Переменная 2
Rueng SDocPageText Importance Konctainte is creationeeme Introduction States Introduction Introduction States Introduction States Introduction States Introduction States Introduction States	Bearpanove Bleates	 переменную длаграммы в выражение Значение 2
Koncraisme is creationermal Hombern Biotechnic Biotechnic Biotechnic Biotechnic Imagenermatic Hombern Imagenermatic Imagenermatic	Ruena. SDocPogeText	Переменная 3
Covernise преобразования Covernise преобразования Covernise преобразования Covernise преобразования Covernise преобразования Covernise преобразования	Koncasha a chagologona	3novenne 1
StratesPyttern	STanies	(Любой тип даннык) Новое вначение переой .
	Strutes-Pattern	The second

Указываем Имя атрибута, которое ранее было задано в Шаблонизаторе, используя кавычки.

2 Crysteen o connecter part - Crasena, pag. Tanegartup sudicione Bullo - Crigaterna	
🗈 🌢 1) Ongara sezzgawił dośn 🌗 2) Objektowa gosywar i 🕞 Rynnegora 🕂 7) "Jefwara styrktyr 🚡 6) Bartowara rozwos cesargeniana	
C @ 44000km100 @ 1000 monored 0 (architect 0) equal 0 monored 0 € € · / 20 threaded 0 (architect 0) equal 0 (b readed 0) (architect 0) (c readed 0	Argelyne Constant & Cons

Далее необходимо задать имя переменной, с которой мы продолжим работать. Имя переменной может быть задано произвольно, например, \$AttributesZakaz или \$MyOrganization.



Присво	ить значение	переменной	
Блок ус или нес переме (числом	танавливает н кольких пере нной может б а, текстовой ст	овые значения для одной менных. Новое значение ыть как константой грокой), так и выражением клах разменных и/или	•
функци	й, В новом зна	ачении переменной	
функци	й, В новом зна Найти	ачении переменной	• ×
функци Функци Пере	й, В новом зна Найти менная 2	ачении переменной \$AttributesZakaz	×

Таблица

Присваивая значение переменной для таблицы (если таблиц несколько, то массив таблиц), необходимо использовать индекс таблицы, который был присвоен в Шаблонизаторе автоматически (по порядку). Вручную присвоить значение переменной невозможно.



На панели свойств выбираем значение переменной, кликаем на «…» и редактируем выражение. В открывшемся окне выбираем категорию *\$Tables*. Добавляем столько таблиц, сколько было добавлено в Шаблонизаторе и указываем индекс каждой таблицы.

Редактировать выражения		(Cashrie
37abies (0) 57abies (1) 57abies (2)		Прискоить значение переменной Блох устанализает новоне эканения для дакой или некольком переменном. Иновая ракиенов переменном исхаят быль как санстватой (маская, техсторой) гак.
-	. The state even of the state o	21 24 (Parties 🗙
Contraction of the local distance of the loc		, Communities
income in the second	 Even diverses of address in the last of t	Переменная 1 🥢 🞽
Caracopite	Дважды цалочная, чтобы аставиты	3+8+8+9+9 1
Distances	Geryte Hermite entering the rect: 30 Stables Gassie sadesore wrote potenties and	Переменная 2
Deserptions	Endex nepewerriyo granpavou a suparative	Значение 2 —
Suena	Dochgelint (Переменная 3
Константы и спецсиналы	SAttributes	Beaumane 1
Полизные преобразование	Allation	(Любой тип данных) новое значение первой переченира. Свойство игнорируеток, если
Математические функции	DirdesPatters	



Далее необходимо задать имя переменной, с которой мы продолжим работать. Имя переменной может быть задано произвольно, например, \$TablesProducts или \$MyTables.

	Основные свойства	×	
[Переменная 1	\$TablesProducts	
	Значение 1	\$Tables [0] \$Tables [1] \$Tabley \$2]	

После того, как Имя и Значение переменных были заданы, можно приступать к дальнейшей работе.

2.2 Работа с многостраничными документами

При создании шаблонов для многостраничных документов необходимо определить уникальные якоря, поиск по которым будет осуществляться только в определенном типе документа. Например, в документе УПД может встречаться слово счет-фактура, соответственно, не нужно использовать это слово в качестве якоря ни для УПД, ни для Счет-фактуры.

При необходимости работы с многостраничным документом, бывает, что один документ находится на нескольких страницах.

Для каждого типа документов необходимо создавать отдельный шаблон (например, 1 — Торг12, 2 – Счет-фактура, 3 – УПД), и указать все типы документов в значениях переменной. При этом тип переменной выбрать List (список).

List (Havanshoe shevenive) - D X Rahens cecilists						
Значение	начение		(O	- Charlettin		
УПД.json			0	Переменные		
TOPF12.json			+	- 🖻		
Счет-фактурајson				Имя	Тип	Начальное значе
			5	Переменная1	Number	0 🖌
			5	Reprisonan]	a de la compañía de la	Редактировать

Далее указываем путь к файлу шаблона:



Редактирость выражен	e (h.		Панези свойств		
Blint_doc	det sesses etc.	i tean di kalena matani per Oria Par	Pacnoseers no and Pacnoseer Jagams Brideo vectoreurs	лону ий документ і	по шаблону
Ederport according			Путь к файлу ша	6n SUite	doc
Категория: Попратители С Диаграмия:	Associate lucification in the second		 Выход Априбуты Таблицы Путь к файлу шабло 	SAttrit STable na*	butes es
Вытод Константы и спецониколы Полезные преобразования Матацальности полиции	Sindes SDucPageText SAttributes	Наконте на завенит, чтобы увидеть его основние	Tepenseneer Hun	Turn	Вачальное значи

Указываем Номер страницы – 1, для того чтобы распознавание начиналось с первой страницы. Также указываем путь к файлу.

нель свойств		
Освойства		
Распознать по шаблону Распознает заданный док Видео инструкция	умент по шаблон	ηy
Д↓ Найти Основные свойства	×	, ×
Construction of the second second second	-	
Путь к файлу*	\$"File.Pdf"	/
Путь к файлу* Номер страницы	\$"File.Pdf" \$IndexPage1	-

Завершение работы робота при возникновении ошибки

По окончании работы по создания шаблона, необходимо прописать дальнейший сценарий его распознавания и вывода результатов.

Если робот не смог произвести распознавание документа, то сценарий может быть остановлен. Для этого в Sherpa Designer добавляем блок «Конец». Также можно добавить блок «Лог» для записи ошибки в лог.



Ошибка распознавания может возникать в случаях, когда ни один из шаблонов не подошел или, когда не осталось документов для распознавания. 64

В случае, если в PDF файле находятся несколько документов, и один из типов документов неизвестен роботу (т.е. на этот тип документа нет шаблона), робот будет пропускать этот документ и будет распознавать лишь те документы, для которых есть шаблоны.

Поиск второго и последующих документов в многостраничном файле

После того, как в многостраничном документе был распознан один из шаблонов, необходимо внести изменения в прописанный индекс (прибавить), чтобы дальнейшее распознавание началось не со следующей страницы, а после найденного документа.

Так как в атрибутах существует параметр CountPage, который соответствует количеству страниц, участвующих в распознавании шаблоном, то для продолжения процесса распознавания нужно добавить этот параметр к индексу.

Таким образом, если в многостраничном документе был распознан первый документ, занимающий несколько страниц, при добавлении параметра CountPage дальнейшее распознавание начнется со страницы, следующей после распознанного документа.

Редактировать выражения	- 🗆 x	Панель свойств
		— 💮 Свойства
Sindex+SAttributes ["CountPage"]		Присвоить значение переменной Блас открытальна чивначил эле окной Г.
	Т	Переменная З SindexPages -
	की	Значение 3 SIndex+\$Attributes ["Co
		Переменная 4

После этого настраиваем сценарий на повторное распознавание документа.





- 3. Примеры
- 3.1 Создание шаблона на примере документа «Справка о стоимости выполненных работ»

В случае, когда одним из действий сценария является распознавание и импорт данных из сканированного документа, необходимо воспользоваться блоками из вкладки Автоматизация OCR, которые находятся на главной панели Дизайнера (Sherpa Designer).



В Дизайнере есть предустановленные шаблоны по следующим видам документов:

- Торг12;
- УПД;
- Счет-фактура;



- Счет на оплату;
- Акт выполненных работ;
- TTH;
- Акт сверки;
- Паспорт;
- СНИЛС;
- НДФЛ;
- CAPCHA.

В случае, когда роботу необходимо распознать информацию из иного вида документа, не указанного во вкладке «Автоматизация OCR», потребуется создать шаблон документа. При этом в сценарий проекта добавляется блок «Распознать по шаблону».



По созданному шаблону в дальнейшем робот будет распознавать и обрабатывать информацию из всех документов данного вида (Справка о стоимости выполненных работ и затрат, Акт об оказании услуг и т.д.).

Для каждого вида документов создается отдельный шаблон.

3.1.1 Предобработка документа

Рассмотрим создание шаблона на примере документа «Справка о стоимости выполненных работ и затрат».

Прежде чем робот приступит к созданию шаблона, документ должен быть распознан роботом и сохранен по указанному пути.

Например, при распознавании документа в формате .pdf в проекте сценария добавляем блок «Получить текст со страницы OCR». В настройках результатов во вкладке «Выход» указываем \$DocPageText



Важно понимать, что в некоторых распознаваемых файлах название столбцов таблиц может не совпадать с названием колонок, которые заданы для вывода результата роботом. Также некоторые столбцы могут отсутствовать вовсе или могут отсутствовать некоторые границы таблиц.

Проверить и сопоставить названия столбцов необходимо в сценарии проекта «Определить колонки.process», они задаются в настройках справа: «Панель свойств» — «Переменные».





Например, после обработки документа «Справки о стоимости работ и затрат» в результате роботом должны выводиться следующие колонки:

Summa	сумма
Price	цена
SummaNDS	сумма с НДС
Stavka	ставка
Name	имя/название
Count	количество

Но при создании шаблона мы видим, что в самом документе часть этих данных отсутствует.

3.1.2 Создание шаблона

Создание шаблона происходит на верхней панели основного меню Дизайнера во вкладке «Разработка» — раздел «Запись» — «Извлечь по шаблону».



Извлечь по шаблону

При нажатии на кнопку «Извлечь по шаблону» открывается обработка «Генератор шаблонов» (Шаблонизатор).

При создании нового шаблона происходит автоматическое распознавание документа. Так как мы создаем шаблон документа, с которым ранее не работали, нам необходимо создать новые якоря и привязки к тем значениям, которые будут обрабатываться роботом именно в данном типе документа. Поэтому на сообщение во всплывающем окне «Вы хотите импортировать все таблицы из документа в шаблон» отвечаем «Нет».

10%	ArpaDyna
Dergra TPA	+ 0; 1 Apresses 3 8 🛧 🔸

Распознавание текста

Распознавание текста в Шаблонизаторе производится с помощью встроенных и внешних модулей распознавания:

- Tesseract OCR
- ABBYY Cloud OCR
- ABBYY Fine Reader
- Microsoft OCR
- YandexVision

В платформе Sherpa интегрированы несколько ОСК -модулей. Два из них – офлайн (поставляются вместе с роботом), и будут работать без подключения к интернету. Это Tesseract OCR и Microsoft OCR.

Yandex Vision и ABBYY OCR – онлайн-модули, т.е. используют функции соответствующих облачных модулей.

ABBYY Fine Reader – это коммерческий офлайн-модуль, использование которого требует отдельной лицензии.

Платформа позволяет настроить работу сценария с распознаванием изображений и переключаться между этими движками в любой момент времени. Для этого необходимо выбрать желаемый движок из списка.





	Pacements react as and	(increase)		
	Пастарнавання такота на задее об наколична долт Поддерживанные форма	учавного графического фай гонал оффикатории и окофо га спобранатнай догу (ред.)	tra illu voxere subper otaci wigjveli GCR jeg. timp	•
	21 Halter			X
	a desenant sectores			
	Canavrag	«Salamar/v		-
	Flyrs w dashoy*			-
	Rectests	angenia		
	Magen	Tenarez OCR		
	Martia pasantema	Insurant 1318		
	Faux normal Stores	Variables Streeger	1.0	_
	Resonances consisters	Nummer OCE		
	Thursday and	Assessor OCH		
	(manual)			

В случае, если распознаваемый документ высокого качества, с высоким расширением, то достаточно использовать бесплатные ОСR-модули.

При обработке документов низкого качества лучше использовать платные решения.

После того, как основные параметры настроек заданы, документ необходимо обработать. Действие запускается нажатием кнопки «Обработать документ».

Waltz Objectures 🗊 🍉 1) Ompere recognesit dallar 🕨 2) Objecture			
👔 🚳 1) Ompere accognesi) dalla 🕨 2) Olgosbora			
	ana accordance · · · · · · · · · · · · · · · · · · ·		
Merurati D0% · D Orotpusco	ne @ Ropa @ Arpetyner @ Tatresar @ Vigneserratines @ OCR finese @ OCR refereare @ OCR Reveal	Recipolice	
		ОСП надать ОСП наракти ОСС (Аланда левного па трана ОСС (Аланда левного па трана Фолу са страностальна лем Фолу са страностальна лем Маракт денного Таблеж располноста и претите	Уколон/Акт 2 Ва 40 40 45 5 5 Тла Тла Тла Тла Тла Тла Тла Тла Тла Тла
		Internation Societ T. Aspellance Internation Internatione Ungroups and internatione	

Обработанный документ выглядит так, как показано на скриншоте выше.

Сохранить шаблон

Для дальнейшей работы с шаблоном его нужно сохранить, нажав на иконку ⁽³⁾в левом верхнем углу основной панели действий. Далее указываем путь для сохранения шаблона и имя файла. Имя файла указываем исходя из вида документа, например, «Справка о стоимости выполненных работ и затрат».

 Ben patjern - Crysnus, gef" - Texepoint patientes Budg - Odpubrina Bu 1) Orspurs inclassed data D) Objedorate goopering 	• 🖸 Appendium: 🕂 1) Apticians attractif. 🍃 () Businesses manage	tanan	
Macurat: 20%	@ Arputyne @ Tativesae @ Herrence rativesa @ OCR torous @ OCR rativ	man @ OCE.freese	No. No.
	Conperante de Conperante de Conperante de Ladoneur, SterpetRA Propagamente + Utator mina Constantino de constante Constantino de constante Constante br>Constante de constante Constante Constante Constante	Artospessae jimaan mars.	тили Чланичина Чи тили 2 тал-макина Раз про-макина Раз Пр
	Kopores Diffuençario Inte quelles Crosesacina suntanesense pation a sergar Ter quelles (porr) Ter quelles (porr) Ter quelles (porr) Coperte name Coperte name		we fuer roomscarce personaleses rects
		山 500 丁 400 田 745 田 745	terne naar naar Talinaa,

3.1.3 Работа с таблицами

После сохранения шаблона переходим к работе с таблицами.




В левой части основной панели настроек выбираем вкладку «Таблицы» — нажимаем кнопку «Импортировать» Э.

Импортируем все таблицы документа, для этого в появившемся окне «Вы хотите импортировать все таблицы документа в шаблон?» выбираем «Да».

На рабочей панели справа отобразятся все таблицы шаблона, которые были распознаны. Их может быть одна-две и более. В нашем случае на скриншоте ниже мы видим, что OCR-модулем распознаны 3 таблицы шаблона.



Каждой таблице автоматически присваивается номер, нумерация начинается с 0. Также отображается количество колонок каждой распознанной таблицы. Именно по этому значению мы можем определить с какой (или с какими) таблицами мы продолжим работать.

В нашем примере нас интересуют данные из таблицы с 6-ю колонками, соответственно, остальные таблицы (с 2 и с 1 колонками) мы можем удалить из окна панели.



Tab	лицы	1					
+	亩		Колонка:	+	亩	ľ.	Ð
D.	Табли	щa		Таб	пица	0 (Колонки 6)	
p.	Tatiny	i LUN	1	Таб	пица	1 (Колонки 2)	
D.	Табли	1110		Таб	пица	2 (Колонки 1)	6

Удалить таблицу можно нажав на значок на верхней панели окна, предварительно выделив таблицу, которую нужно удалить.

Индекс

Индекс OCR-таблицы. В случае, если Шаблонизатор извлекает данные лишь из одной таблицы сканированного документа, необходимо указывать индекс, равный 0. При этом для улучшения привязки рекомендуется указывать минимальное количество колонок.

Таблицы	
+ 📋 🕏 🏂 Колонка: 🕂 📺 🏒	é Đ
🔺 Таблица 🥒	Таблица О (Колонки
Индекс	0
Пропустить строки	1
Строк в заголовке	1
Связующий якорь	
Минимальное количество колонок	5
Максимальное количество колонок	0
Колонки	Колонки (б)
Рабочие данные	Данные, вычисляемые

Минимальное количество колонок

Указывается минимальное количество колонок в ОСR-таблице. Все таблицы с количеством меньше указанного будут пропущены.

Максимальное количество колонок

Указывается максимальное количество колонок в ОСR-таблице. Все таблицы с количеством больше указанного будут пропущены.



🕂 💼 🌲 🏚 Колонка: 🕂 💼 🗌	Ð
 Табляца 	Таблица 0 (Колонки
Индекс	0
Пропустить строки	1
Строк в заголовке	1
Связующий якорь	~
Минимальное количество колонок	0
Максимальное количество колонок	5
Колонки	Колонки (б)
Рабочие данные	Данные, вычисляемые

При создании шаблона «Справки о стоимости выполненных работ и затрат» указываем минимальное количество колонок равное 5, т.к. нас интересуют данные из таблицы с 6 колонками. Остальные таблицы данного документа содержат менее 5 колонок.

Таким образом Шаблонизатор продолжит работу с интересующей нас таблицей.

При создании шаблонов для других типов документов (с большим количеством таблиц и колонок в таблицах) можно указывать и максимальное и минимальное количество колонок, тем самым задавая интервал, в который будут попадать таблицы с интересующими нас данными.

Проверить

Для проверки корректности отображения табличной части, из которой будут импортироваться данные нажимаем кнопку «Проверить» на верхней рабочей панели.



Также проверку можно запустить нажатием клавиши F5.

Пропустить строки

Указываем количество строк, которые нужно пропустить для того, чтобы данные этих строк не попали в результаты работы робота.

При выборе значения, равного 1, — первые строки будут пропущены. Наименования колонок будут соответствовать значениям первой строки.

Если задать названия колонок в шаблоне, то эти значения не будут изменены.

В нашем примере необходимо пропустить 5 верхних строк таблицы, т.к. значения интересующих нас данных указаны в таблице документа начиная с 6 строки.

Для этого в строке «Пропустить строки» указываем число 5.



Таблицы				
🕂 📋 🌲 🏚 Колонка:	+ 💼 🛛 🕀			
 Таблица 	Таблица 0 (Колонки 6) 🦯			
Индекс	0			
Пропустить строки	5			
Строк в заголовке	2			
Связующий якорь				
Минимальное количест	5			
Максимальное количес	0			
> Колонки	Колонки (б)			
Рабочие данные	Данные, вычисляемые в процессе (

Строк в заголовке

При построении таблицы, заданное количество строк будет использовано для создания заголовков колонок. При указании значения 0 заголовки будут взяты из шаблона колонок, в противном случае название будет суммироваться из ячеек колонки по заданному количеству строк.

Например, в шаблоне «Справки о стоимости выполненных работ и затрат» необходимо указать количество строк заголовке равное 2, т.к. в заголовок должна попасть информация, размещенная в двух верхних строках таблицы.



После выполненных действий необходимо проверить промежуточный результат. Нажимаем «Проверить» — далее «Выполнить полное сканирование».

Результат полного сканирования



Результат полного сканирования отобразится в виде таблицы с

импортированными данными из табличной части сканированного документа.



Определить колонки

На этом этапе часть данных может отражаться некорректно, в частности, названия колонок могут не совпадать с теми значениями, которые ранее были заданы в сценарии проекта «Определить колонки.process» — в настройках «Панель свойств» — «Переменные».

Для корректировки выводимых результатов необходимо вернуться в сценарий проекта «Определить колонки.process», сопоставить значения переменных, заданных в настройках «Панель свойств», добавить указанные здесь колонки и определить для них названия.

1	Dephenosy denia: Parena Mannani, Tanya a		New York Contraction	There and the Designed State
Conservation and the second seco		Respective Continues Intern Continues	O Property (Chromate and and a second s	2011 inter-

Добавление колонок «Количество», «НДС Сумма», «Ставка».

В нашем примере в результатах полного сканирования отсутствует несколько колонок, в том числе колонка «Количество».

Название колонки должно соответствовать значению, присвоенному данной переменной «количество».

T	🖞 🛢 🗿 Колонки	+ 🗊 😑
	adhaug	Tationa 0 (Konoreor 7)
	Hydenic.	D
	Пропустить строки	5
	Crock a sanonoske	2
	Святующий якорь	
	Миничальное количес	5
	Максимальное холиче	1 P
	и Колонки	Kondeiser (7)
	Колонки	0 : Ees versen 1
	 Konowsa 	1 : Ees weenen2
	Konowa	2 that we wanted
	NO.ROHKIE	3 . beo www.enwit
	Konneka	4 Data weapons
	A Konturna	5 - Kommerson
1.00	a ponovela	
	Visco .	Количество
-	Twn gammaix	Test
	Рабсчи данных	Данные вычественые в процессе
14.14	1001 -	
Hear Hear	nec so xonoren n OCR tatinuy konorensi	в, не которуе ссылантся табляця с
Mean Mean Mean Mean Mean Mean Mean Mean	нис на колонии в ОСЯ пиблиц солонкой Настройки	е, не которус ссылантся таблица с
	нис колонног в ОСЯ теблиц колонной Настройки Пагра	, не которос соытается табляца с
Huan Hean Satalia Satalia T	nec concentration of DCR radional concentration Hacipolika Respo Argedyne	, не которор соытантся табляца с
	nec ex source-rear is DCR technologi exercised Histopolites Respo Arconform Todowas	, не которие соытантся таблица с



Tationage 0	ects et	OK ONE M			1
не свання тускавых контлексая, их, объекта, яклая выполнение дабот, рудование, заграт	Kea	Стоинестьс начала проведения рабет	выполненных. руба нечеля года	работ и запратия три часле за отнетный переса	Колчество
накомя работы по сносу подсебных пенящений		108-008	100 000	100 000	
ва невосоннатных перегородок		60 000	60 000	60 000	
e (1044-00/16 (1008-0316)=0.00		120.000	120.000	120 000	
			Vitara	280 880	
	од разии ОС Обесенте Т Ла Таблица: 0 на обеста, ечала сонтанская, по биста, сонтанска	од разова Собесения. Т Перери со Табонци, 0 на оденка пускавка контанска. на объекта, какае контанска. на объекта, какае контанска. какае рабони по осор, тодобъел понещиней ка неколнатика прогладок	ад ражила Собектити Т. Полери всох околе Табликци: 0 мнобликотик, макая на политическая, ин объектик, макая на политическая, ин объектик, макая на политическая, удожение, затрат чтакочна работы то соксу годотбыля потичешиена на политическая на порегородская и понициона на понициона на порегородская и понициона на порегородская и понициона на порегородская и понициона на	од разака • (Constraint Constr

Добавление недостающих колонок производится на правой рабочей панели Шаблонизатора во вкладке «Таблицы». Нажимаем на кнопку +, добавляем колонку, присваиваем колонке имя «Количество».

При повторной проверке в результатах полного сканирования Шаблонизатора появится колонка «Количество».

Производим проверку остальных колонок и корректируем их названия в зависимости от значений переменных в сценарии проекта «Определить колонки.process».

При создании шаблона «Справка о стоимости выполненных работ и затрат» аналогичным образом добавляем колонки: «НДС Сумма», «Ставка».

В случае, если название колонки не соответствует названию переменной, но содержание колонки соответствует значению переменной, то создавать дополнительную колонку не нужно, достаточно изменить название колонки в Шаблонизаторе.



Taö	лиц	tы					
+	đ	**	Колонка: +	亩	Ð		
	Таблица Индекс Пропустить строки Строк в заголовке			Таблица 0 (Колонки 7)			
					999 020 99 197 0 V		
				2			
		Связующий	якорь				
		Минимально	е количест 5	5			
		Максимальн	ое количес 0				
		Колонки	Ko	тоннки (7)			
		Колонка	0 :	Без име	ни 1		
		 Колонка 	1:	Без име	ни2		
		Колонка	2:	Без име	ниЗ		
		 Колонка 	3:	Цена	1		
		Инд	екс 3		K		
		Имя	Це	18			
		larn	данных Тех	4			
		» Колонка	4 :	Без име	ни5		
		» Колонка	5 :	Стоимос	ть		
		Колонка	6 :	Количес	тво		
	D	Рабочне дан	ные Да	HILLIO, BUNK	сляемые в процессе		

ехини работыс Просто	é pexose · O Odvetaire T Resept	ACELONON
ites Seamon	Tetroux 0 *	-
	не-свание турсавыя сонтоннойя. разбъектов, нохов высосноеные работ. узаконные ватрат	Стонностьс начиля проведения работ
	-reasoning patients no energy registerious mentionents	100 000
	an exercision press reperiptions:	60 (000
	a crownochu crootherepwance	120 000

Например, в таблице с результатами Шаблонизатора название колонки «Стоимость начала проведения работ», а название переменной «Цена», тогда мы вносим изменения в название колонки, предварительно убедившись в том, что именно в этой колонке отображается цена.

Аналогичным образом необходимо изменить названия всех колонок, которые не совпадают с названиями переменных.

Нумерация колонок таблицы в Шаблонизаторе начинается с 0.



Таб	ілис	ты				
+	Ē		Колонка:	+	۵.	Ð
•	Tad	блица		Taő	блица О (Колон	ки 7)
		Индекс		0		
	Пропустить строки			5		
		Строк в заг	оловке	2		
		Связующий	якорь			
		Минимальн	ое количест	5	1	
		Максималь	ное количес	0	-	
		Колонки		Коло	очки (7)	
		⇒ Колонка	a	0:1	Без имени 1	
		» Колонка	a	1:1	Без имени2	
		Колонис	a	2:1	Без имениЗ	
		> Колонка	а	3:1	Цена	
		 Колонес 	a	4 -1	Без имени5	
		 Колонка 	a	5:0	Стоимость	
		 Колонко 	a	6:1	Количество	
	5	Рабочие дая	ные	Дан	ные, вычисляем	ые в процессе

В случае если название колонки сохранено по умолчанию и содержит в себе числовое значение (например, «Без имени1»), то в таблице результатов Шаблонизатора название указанной колонки определится автоматически и будет содержать в себе текст из соответствующей колонки сканированного документа.

Если название колонки мы изменили вручную, то в таблице результатов Шаблонизатора будет отображаться именно заданное нами название колонки.

Поэтому, рекомендуем обратить особое внимание на часть табличного документа, в котором название колонки указано в объединенных ячейках. Даже если при создании шаблона документа название колонки из объединенных ячеек было распознано корректно, в других документах название может быть указано со смещением.

В нашем примере слово «Стоимость» располагается в 4 ячейке таблицы, но в некоторых документах может сместиться в 5 ячейку. В таких случаях рекомендуем переименовать все ячейки, названия которых могут быть смещены.

		Номер докум	ana	Дата сос		Отчетвай период
	СПРАВКА	14		31.01	2016 03	06.2016 31.08.201
	о стоимости выполн	ЕННЫХ РАБО	TR3	ATPAT	50.00	
He-	***************************************			Стоньюсть	DARDO OFFICIAL ST	perior at tampat,
Meg	MANDACHINE STREET, STATE STATE STATE	man			py6.	
по по- рязку	объекток, килок выполненных работ, оборудования, затрат	Ken	ab	начала оведенися работ	с начала года	в том числе за отчетный первод
1	2	3	1	4	3	6
	Всяго работ и затрат, папиочаныват в стопность набот					



По завершению работы с таблицей в Шаблонизаторе необходимо произвести проверку и удостовериться в корректности результатов полного сканирования.

3.1.4 Якоря

Точки на сканированном документе, с помощью которых робот определяет, в какой части документа размещаются определенные данные.

При помощи якорей мы задаем границы определенных атрибутов. Якорей может быть сколь угодно много, в зависимости от типа документа.



В нашем примере необходимо получить данные следующих атрибутов:

- «Инвестор»
- «Заказчик (Генподрядчик)»
- «Номер документа»
- «Дата»

Создание якоря (атрибут «Инвестор»)

Рассмотрим процесс создания якоря на примере атрибута «Инвестор».

Справка о стоимостијаоп - Справка, до	ат - Генератор шиблонов	
@ails Ofpationsa	/	
🕄 🕒 1) Orepurts secregatail della 🗼 2	1 Обработать документ 🔹 💽 Проверять 🖙 3) Доберать апребут 🚔 4) Вытолного полное скотородонее	/
- 0 0	Ordenment Officer Officer Officer Officere officer Officer	
- p p	and and a which a stand a show must a comment of comments	
	The second se	+ · · · · · · · · · · · · · · · · · · ·
	Casada Anna anna anna anna anna anna anna a	# Rops 1
	Participation of the second seco	10 1
	Page as (0.7.2 (0.00)	Люрные привязки Привязки (2)
	Service 602 Service Contract Contraction of State (1997)	Tan Text
	Taxing Precision (Inc. Heat Inc.	Seere
	Second Diff. And Diff.	Pexer pational Default
	THERE AND A THE ADDRESS OF THE ADDRESS AND ADDRESS br>ADDRESS ADDRESS ADD	Pesannuru-e crose True
	Carifia Executed addite del Assesser an 1920 Charina	Kennerme Enorce 0
	Enclosed and a company of the second s	Oncase
		Рабоция данные Данные видассваеные в горциос
	Sciences and Scien	Conventence acces Second (0)
		Contraction and a state of
	Reary projections Jacob Strategies 1	
	KORANG LAND LAND LAND LAND LAND	
	D-CORRECT: BUILDED BEITHER STATE FALSE FALSE FALSE	Buips
	Re- Conserved annumers of the strape.	1
	100 Discourse systems and the second states and pro-	
	ar so- advances, and a successing pales. The cannot be the weat to	
	parties and partie	
	Description sugar	
	Contracting period	
	anomasa principa any manima.	
	10.00 10.00 10.00	
	disas remaining disastructure da disastru da disastructure da disastructure da disastructure da disastructure da disastructure da disastructure da disast	
	key served 194	A Lacitorei 🦌
		A Deces
	Canada Michael Brance Control	
		T Applying
	88	
	Constant In Acres From From 75	III Totivusv
	1000 Historia	The strength of the strength o
		III Album again

Для создания якоря переходим во вкладку «Якоря» в правой рабочей области Шаблонизатора и нажимаем кнопку + в верхней части панели. В этот момент становятся активны «OCR блоки» и «OCR таблицы» для удобства выбора определенного блока в качестве якоря.

Для того, чтобы Шаблонизатор мог корректно определить местоположение атрибута в документе, необходимо ограничить его границы. Ограничение верхней границы задается по тексту, находящемуся на строку выше самого атрибута. В нашем случае это блок «Форма по».

Для присвоения значения якорю достаточно кликнуть левой клавишей мыши по OCR-блоку (слову), которое мы выбрали в качестве якоря. Также можно указать значение вручную, написав якорное слово в строке «Значение».

В случае, если мы укажем в качестве якоря слово «Форма», то якорь будет установлен неверно, т.к. в документе есть несколько ОСК-блоков, распознавших это слово.

		+ • • •	Npomassan 🧔 🏚 🛧 🔟
Townson & Contractory	Transmiss PL Lat	# Riops	1 Форма
Paper at 1272	646. 182200	ID Rospiske rep	пизки Пранаярии
hearing 000 Theorem (Marcen Epicongroup part) 0 0000	theme:"	Ten	Test
Annual Temportal Dif Teach (Lines Teachers (1971) (1971) with TEAL	ADM T	Fexam pelo	ne Default
And Annual Contraction of the Annual Contract	104 10	Pasaeno-roie	crose True
manual and has preventioned in (4.22)		Dracase	Snewcow D
	a f a fine	Petowe and	ные Данные, вычисляемые в процессе
pro maisman ["		Поделение	в неори Якорик (D)

Поэтому меняем указанное значение якоря на «Форма по». После этого необходимо произвести проверку, чтобы убедиться, что якорное слово определилось корректно.

Когда значение якоря задается нажатием клавишей мыши на ОСR-блок, к значению добавляется знак «*», если значение указывается вручную, добавлять этот знак необязательно. При использовании якорных привязок знак «*» нужно удалить и после этого добавить привязку.

В случае, если значение содержит два и более слова, — между ними необходимо ставить пробел, т.к. он выступает разделителем слов.



Далее добавляем якорь для атрибута «Инвестор». Для корректного распознавания информации об Инвесторе, необходимо ограничить якорными точками данный атрибут слева и справа.

Слева якорем будет выступать слово «Инвестор». Добавляем якорь, нажимаем левой клавишей мыши на ОСК-блок «Инвестор» и Шаблонизатор автоматически заполняет строку «Значение», тем самым определяя якорную точку.

1.12 Sector State State State State	+ · □]
Conjunction of the Alice Andrew State Andrew	Родь Родь Лорь Лорь Лорь
Taxanan ("managaman") The second and the	Districtione surface Top" Peacer participations Default Pasagenariate cross True
Bas assesses or to (5.2.5) Derived togetic longer (1) 1000 100 100 100 100 100 100 100 100	Количество блоков 0 Описания и Рабоже данные Данные, вычисяленые в процес и Оссаниения посто — Якора (00

Справа якорем будет выступать слово «по». По аналогии добавляем якорь в рабочей области Шаблонизатора, нажимаем на ОСК-блок «по» и производим проверку.

После проверки видим, что якорь установлен неверно, т.к. в документе есть несколько OCR-блоков, распознавших слово «по».

Для корректного определения якоря осуществляем якорную привязку к якорю «Форма по», который находится строкой выше того блока, который мы хотим указать в качестве якоря. Для этого нажимаем на кнопку «Добавить привязку», затем кликаем по якорю «Форма по», тем самым привязываем наш



новый якорь к OCR-блоку, который располагается строкой выше нужного нам

якоря.

Feits Occutions		Construction of the local distance of the lo
🗄 🕼 1) Ongaria wangawél dala 🗼 2) Olipatinana googwaar 🔹 💽 Aparapana 🕂 🔄 Aphanina opolys 🚔 4) Batanana repolys		
Macureli 19% · P P Orošpanana: @ Angelijita: @ Talenaja: @ Hightever rationaj @ OCR factor @ OCR rationaja: @ OOR Ansee	Aven	· · · · · · · · · · · · · · · · · · ·
	+ ·	enter 2 2 4 4
	 Ropsi Ropsi Ropsi Ropsi Ropsi Termina Ropsing Ropsin	1 Popear no 2 mesor rop* 3 mo* 1 5 motor (% 1 : fotom % fixes Tat 1 : fotom % fixes Tat 1 : fotom % fixes Tat 1 : fotom % fixes

Для того, чтобы Шаблонизатор определил координаты создаваемого якоря относительно якоря «Форма по» указываем параметр привязки «Bottom». Этот параметр говорит о том, что создаваемый якорь будет привязан к якорю «Форма по» и находится строкой ниже.

 Якорные Прив 	привязки язка D якоря	Привязки (1) 0 : Bottom > None 0
ſ	Тараметр	Bottom
C	бъект привязи	None
Г	Іринудительно	False
C	мещение Х	0
C	мещение Ү	0

Создание якоря (атрибут «Заказчик»)

По аналогии создаем якоря для атрибута «Заказчик». Слева якорем будет выступать слово «Заказчик». Добавляем якорь, нажимаем левой клавишей мыши на OCR-блок «Заказчик» и Шаблонизатор автоматически заполняет строку «Значение», определяя якорную точку.

Справа якорем будет выступать слово «по». По аналогии добавляем якорь в рабочей области Шаблонизатора, нажимаем на ОСR-блок «по», производим проверку и добавляем якорную привязку. В этом случае привязку осуществляем к якорному слову «по», которое располагается строкой выше в документе. Указываем параметр привязки «Bottom» и производим проверку.

Странна о стоямасти дант - Стронна, д	rg - (merpetate set-more	
Cuite Odpedoma		
📳 🕒 I) Orepurte erospesiek dalle 🕨	2) Objesteraria geogramier 🔹 🕞 Opposegania 📫 3) (Antoneria erpetiyn 👌 4) Banneriaeria norace conseponance	
Menurat: 30% - 8 8	Oroópasana: @ Respe @ Arperbyne @ Tatinese @ Heprese ratinese @ OCK tensor @ OCK retineses @ OCK Research	Recent
		Araps A
	Separa segura has beauty	Information and Annual
	Andre	🖩 Uspran salas.
lane farmer dire	and the second sec	
Citational Society Rypon	a Conache	

Атрибут «Заказчик» необходимо ограничить не только справа и слева, но и сверху. Для этого в качестве якоря выбираем слово «организация». Создаем еще один якорь, редактируем название так, чтобы с обеих сторон слова стоял знак «*» – «*организация*». Проверяем корректность расположения якорей.



При создании якорей, в значении которых содержится несколько слов, необходимо указывать знак «*» в конце каждого слова.

Если сочетание слов встречается в документе единожды, то второе слово можно написать сокращенно. Например, в качестве якоря по номеру документа можем указать «номер* докум*».

Crystel a crossette jusi - Crystel agg - Developing galdering	A CONTRACTOR OF
taks Objections	
🔯 🕒 1) Ongarna sozijanal djala 🕨 7) Olgabistan gangaren - 🗉 Ripoletjans 🕂 3) Zafanon Aparija 📚 4) European sozinet comoptense	
Macarale 1971 - P P Ovolganama @ Bates @ Agentyna @ Tatanam @ Henrew Infana, @ OCR farmer @ OCR referance @ OC	anan Pasys
	🛉 🕂 • 🖻 🖞 Tyrecour 🌲 🏚 🕈 🗣
	Press P
DETERMINE THE ADDRESS OF ADDRESS	Patuditu-Fulle circleal Trum Konstructure disexces 3



Если в начале значения якоря присутствует лишний символ (чаще всего это скобка), мы можем удалить этот символ и вместо него поставить знак «*». Тогда получим значение **организация** вместо (*организация**.

Завершающий якорь

При создании шаблона необходимо указать якорь, который будет завершать документ. Т.е. по этому якорю мы можем определить, что в документе таблиц больше нет и работу с таблицами можно завершать.

В нашем примере завершающим якорем будет слово расшифровка.



В нашем примере завершающим якорем будет слово расшифровка*.

Шаблонизатор позволяет задать для якоря несколько альтернативных вариантов, по которым робот будет распознавать завершающий атрибут.

Добавим слово **подпис**^{*}, укажем намеренно без окончания, т.к. в документе может быть указано как подпись, там и подписи.

Если в документе отсутствует слово *расшифровка**, то робот будет искать слово *подпис**

Для завершающего якоря указываем Режим работы – LastPage.

3.1.5 Создание атрибутов

Следующий этап в создании шаблона - создание атрибутов.

Атрибуты – позволяют определить и обозначить области документа,

данные из которых будут распознаваться роботом и импортироваться в таблицу «Результаты».

Атрибуты «Инвестор» и «Заказчик»





В нашем примере в таблицу с результатами должны импортироваться данные об Инвесторе, Заказчике, а также Дата и Номер документа

В Шаблонизаторе необходимо добавить эти атрибуты на рабочей панели справа: выбираем «Атрибуты», нажимаем кнопку добавить, присваиваем имя атрибута «Инвестор».

Для корректного распознавания области документа, данные из которой будут импортироваться в результаты, для атрибута необходимо осуществить привязку к ранее созданным якорям.

The second secon	а Атрибут Иносотор Ина Иносотор
	Programme Topenance (2) Programme Topenance (2) Programme 2: Right -> None Topenance 3: Left -> None
aparate (1) Anagaman (10) (1) Spanned (1) (1) (1) (1) (1) (1) (1) (1) (1) (1)	Discope 3 Tapawerp Lett Other course None
An average segment and a segme	Determined Table Oversitered X D
Proceedings and the second sec	Oversenve Y 0

В нашем примере якорем справа будет выступать якорь «Инвестор», слева якорь «по».

Якорные привязки к атрибуту «Инвестор»

Для создания привязок в созданном атрибуте нажимаем «Якорные привязки», далее «Добавить привязку», затем левой клавишей мыши кликаем на 89

якорное слово «Инвестор». Аналогичным образом создаем привязку к якорному слову «по». Производим проверку, в результате которой область данных нашего атрибута должны подсветиться синей рамкой, как на скриншоте выше.

Чтобы ограничить область данных атрибута сверху, создаем третью якорную привязку к OCR-блоку «Форма по». Нажимаем «Добавить привязку», кликаем левой клавишей мыши на созданный ранее якорь «Форма по». В параметрах привязки выбираем значение «Bottom», т.к. в качестве ориентира мы будем использовать нижнюю границу данного якоря.



Помимо параметра в привязке для данного атрибута необходимо указать «Объект привязки». Т.к. данные атрибута могут располагаться выше строки, которую мы выделили при помощи первых двух якорей слева и справа.

В строке «Объект привязки» указываем «TopEdge» – верхняя грань.

Crasses o consisting on 1 Crasses pay - Foreporte advises	- 41
Nation Dégadiores	
📚 2) Oroparts recordenal dester 🕨 2) Objectionens gesynner 🔹 💽 Appendents 🕂 2) Aptientes etpedyr 🚔 4) Bastonents roomste conseponaese	
aunst: 10% · P P Orotposanc @ Rope @ Argedyne @ Tativesa @ Vegresseratives @ OCE hours @ OCE Interesa @ OCE. Inves	на Анрибути
	4 C 2 Fipennes 2 2 4 4
	Apply: Next Type: Ne

Якорные привязки к атрибуту «Заказчик»

По аналогии создаем атрибут «Заказчик», устанавливаем привязки к ранее созданным якорям «Заказчик» справа, «по» слева и «организация» сверху. Для привязки к якорю сверху устанавливаем параметр «Bottom», «Объект привязки» указываем «TopEdge».



В результате по каждому атрибуту должно быть указано по три якорные привязки: справа, слева и сверху.

После проверки в шаблоне каждый атрибут будет отображаться в виде синей рамки. В случае, если границы рамки не захватывают всю область данных атрибута, необходимо внести изменения в настройки.



Ограничение линиями

В некоторых документах названия полей, которые мы используем в качестве якорей могут располагаться не на нижней линии с описанием поля, а сверху или посередине, как на рисунке ниже.

manual in Astart	SUBJECT.
THE ALL OF THE	PERSON
NAL ORTO	NHHH
NL OKTO	66200000
es 0070	
rtpast selocep	17 42 203
	0000000000072

В таких случаях для корректной привязки рекомендуем воспользоваться дополнительным параметром «Ограничение линиями». Предварительно нужно проверить присутствуют ли в шаблоне ОСК-линии, ограничивающие созданные нами атрибуты.

Нажимаем кнопку OCR-линии на верхней рабочей панели и проверяем наличие линий.



Линии подсвечены зеленым цветом, и мы можем использовать их в качестве дополнительного ориентира в работе с атрибутами.

На верхней рабочей панели выбирает «Атрибуты», кликаем левой клавишей мыши на один из атрибутов, после чего он становится активным и в правой рабочей панели открываются инструменты для работы с выбранным атрибутом.



Выбираем «Ограничение линиями», из выпадающего списка выбираем значение «Down». Таким обратом мы ограничиваем нижнюю границу атрибута, которая теперь будет ориентироваться на ОСК-линии.

Обязательно проверяем корректность отображения границ атрибута. При необходимости меняем значения данного параметра.

Создание атрибутов «Номер» и «Дата». Якорные привязки



Атр	ибу	ты				
+	R	1	Привязка: 📚	\$	٠	*
p.	Атр	мбут		Заказ	зчик	
p.,	Атр	ибут		Инвес	стор	1
4	Am	ибут		Номе	PK	(
VMR		Номер	7			
		Икорн	ые привязки	Привязки (1)		
		4 N	 Привязка 		7 : Bottom /> BottomEdge	
ID ski	ID якоря	7	×			
			Параметр	Bottom		
			Объект привязи	Bottom	Edge	
			Принудительно	False		
			Смещение Х	0		
			Смещение Ү	0		

Добавляем в правой рабочей панели атрибут, задаем имя атрибута *Номер*. Добавляем привязки к ранее созданным якорям.

В качестве привязки слева используем якорь *Справка**, в качестве параметра выбираем *Bottom*, также указываем *Объект привязки – BottomEdge*.



Создаем вторую привязку к якорю *номер* докум**. В качестве параметра выбираем *BottomLeft* (нижний левый угол). Объект привязки указываем *TopLeft* (верхний левый угол).

Таким же образом производим привязку атрибута к правому углу. Создаем третью привязку также к якорю *номер* докум**. В качестве параметра выбираем *BottomRight* (нижний правый угол). Объект привязки указываем *TopRight* (верхний правый угол).

Таким образом, для атрибута Номер должны быть созданы три якорные привязки.

93



После проверки в шаблоне должен появиться синий прямоугольник, внутри которого отображены данные нашего атрибута.



Создание атрибута Дата.

Атр	мбу	ты			
+	Ę	🖞 📋 Привязка: 🌲	∻ ↑ ↓		
Þ.	ATE	мбут	Заказчик		
Þ	An	мбут	Инвестор		
D.	An	мбут	Номер		
• Aŋ	Att	жбут	Дата		
	11.2.1	Имя	Дата Привязки (3) 7 : Bottom -> BottomEdge		
		Якорные привязки			
		Привязка			
		Бривязка	9 : BottomLeft -> TopLeft		
		 Привязка 	9 : BottomRight -> TopRight		
		ID якоря	9		
		Параметр	BottomRight		
		Объект привяза	TopRight		
		Принудительно	False		
		Смещение Х	0		
		Смещение Ү	0		
		-			

Добавляем в правой рабочей панели атрибут, задаем имя атрибута *Дата*. Добавляем привязки к ранее созданным якорям.

В качестве привязки слева используем якорь *Справка**, в качестве параметра выбираем *Bottom*, также указываем *Объект привязки – BottomEdge*.

Создаем вторую привязку к якорю дата* *состав**. В качестве параметра выбираем *BottomLeft* (нижний левый угол). Объект привязки указываем *TopLeft* (верхний левый угол).

Таким же образом производим привязку атрибута к правому углу. Создаем третью привязку также к якорю дата* *состав**. В качестве параметра выбираем *BottomRight* (нижний правый угол). Объект привязки указываем *TopRight* (верхний правый угол).

В результате для атрибута *Номер* должны быть созданы три якорные привязки.



После проверки в шаблоне должен появиться синий прямоугольник, внутри которого отображены данные нашего атрибута.



Тип данных – DateTime.

Формат значения – вручную указываем формат, необходимый для атрибута. В нашем примере это *dd.MM.yyyy*



В случае, если в каком-либо документе дата документа будет указана как 30 *ноября* 2022*г.*, в результате эта дата будет преобразована в 30.11.2022.

3.2 Создание шаблона с фиксированными атрибутами на примере документа «Счет на оплату»

Фиксированные атрибуты имеют положение и размеры, привязанные только к сторонам документа (масштабу), но не привязанные к конкретным якорям, поэтому их можно свободно перемещать по документу. Фиксированные атрибуты допустимо использовать только если во всех возможных образцах документа необходимые для извлечения данные находятся примерно на одном и том же месте и не смещаются по документу.

При создании шаблона «Счет на оплату», Шаблонизатор распознает все таблицы, которые содержатся в документе. Эта информация отображается на вкладке «Таблицы».

All Ofgederate B O Ofgederate B D Offgederate	чегузоп - Счет на оплату png - Генератор шаблонов						- 0
10 Orspans scooped (44) 10 Ofspedorate gargeers 10 Ofspedorate gargeers	n Otipationia						
UNI Image: Comparation of the property	🖕 1) Отерыть исходный файл 🗼 20 Обработаль дляр	мант • Проверить -	- 3) Добене	ni angenfiya 🔁 🗟 🖬 🖬	normal produced	сканирование	
Butterstell Chold a gaveen portsta scales are depresented to canada in trades a trades a trades a trades a trades a trade of participal potentia a trade potentia trade of participal potentia a trade of participal pote	1015 · @ Orodyamatur @	leope @ Arputhynu @ Tatle	wanger @ 14	о 🕞 инт на табат на атца	R Gauss (1) 00	Ritationages Tathonias	
Between Data approprint to approprint to approprint the approprint to	20441 C.C.S. 25-8-66 Children Control Control						
Contraction of provide viel provide viel provide ratio to particip provide ratio (provide rati (provide ratio (provide ratio (provide rati (Resident Covaria menors many many	Contraction of the local diversion of the local diversion of the local diversion of the local diversion of the	reasons Vice	Alexander 14 contractor	and the second second	3 1000	Таблица 0 (Колония 3
	противном разуная на тарантируется нали	противнова случая не тарантаруется налиние токара на склада. Токар стурскется со факту прилага денет на ріс					
Image: market in the second in the	TAO TANT Allower	Inder	Course of the	0000			
Name One Operation OCO ⁺ "Kownew" Internet Internet Internet OCO ⁺ "Kownew" Internet Internet Internet Internet OCO ⁺ "Kownew" Internet Internet Internet Internet OCO ⁺ "Kownew" Internet Internet Internet Internet Internet OCO ⁺ "Kownew" Internet Internet Internet Internet Internet OCO ⁺ "Kownew" Internet Int	SHOT BOARD - FINGER	Carter, Ale		0000			
Internation Description Description Description DOD "Rearrange" Internation Description Description Description DOD "Rearrange" Internation Description Description Description DOD "Rearrange" DOD "Rearrange" Description Description Description Dot The Annual Million DOD "Rearrange" DOD "Rearrange" Description Description <t< td=""><td></td><td>CV. NF</td><td></td><td></td><td></td><td></td><td></td></t<>		CV. NF					
Construint Cynex 1, Kanners 2, Marting Development Construint Construint Doct and under Ne 10 of 01 despans 2018 r. Construint Construint Doct Toosynamen*, With 000000000, KTII 00000000, 125005, Mockass r. Teepokas yrt, goon Ne 5 Construint DOC Toosynamen*, With 000000000, KTII 00000000, 1100 F, Mockass r. Teepokas, yrt, goon Ne 5 Construint DOC Toosynamen*, With 000000000, KTII 00000000, 1100 F, Mockass r. Husself Apfiel yrt, goon Ne 5 Construint DOC Toosynamen*, With 000000000, KTII 00000000, 1100 F, Mockass r. Husself Apfiel yrt, goon Ne 5 Construint Doct Toosynamen*, With 000000000, KTII 000000000, 1100 F, Mockass r. Husself Apfiel yrt, goon Ne 5 Construint Doct Toosynamen*, With 000000000, KTII 000000000, 1100 F, Mockass r. Husself Apfiel yrt, goon Ne 5 Construint Doct Toosynamen*, With 000000000, KTII 000000000, 1100 F, Mockass r. Husself Apfiel yrt, goon Ne 5 Construint Doct Toosynamen*, Doct Ne 7, yr. Yrt Ne 7, Yrt Ne 7, Yrt Ne 7, Soon 2, So	000 %	U.S. 18		L2		Tatimena II I	
Creating OCC "Rownesses", MHH 0000000000, KTITI 900000000, f25000, Moxxaa r, Teepxaa yr, gene Ne 5 Tocynames OCC "Rownesses", MHH 0000000000, KTITI 900000000, f25000, Moxxaa r, Teepxaa yr, gene Ne 5 Tocynames DCC "Rownesses", MHH 0000000000, KTITI 900000000, f25000, Moxxaa r, Teepxaa yr, gene Ne 5 Tocynames DCC "Rownesses", MHH 0000000000, KTITI 900000000, f25000, Moxxaa r, Teepxaa yr, gene Ne 5 Tocynames DCC "Rownesses", MHH 0000000000, KTITI 900000000, f19607, Moxxaa r, Hussaid Apfeer yn Inschere DCC "Rownesses", MHH 0000000000, KTITI 900000000, f19607, Moxxaa r, Hussaid Apfeer yn Inschere DCC "Rownesses", WHH 000000000, KTITI 900000000, f19607, Moxxaa r, Hussaid Apfeer yn Inschere DCC "Rownesses", WHH 000000000, f25000, Mox and r, Hussaid Apfeer yn Inschere DCC "Rownesses", WHH 000000000, f25000, Mox and r, Hussaid Apfeer yn Inschere DCC "Rownesses", WHH 000000000, f25000, Mox and r, Hussaid Apfeer yn Inschere DCC "Rownesses", WHH 000000000, MOX and r, Worker B, Gong S,	COO MARCHINE					Cepesa 1, Konovea 2	
Cuor Ha onnary Ne 10 of 01 deepans 2018 r. Doctranuum: OOC "Kownewer", WHH 000000000, KTIN 00000000, 125005, Mockaas r, Teepckas yn, grow Ns 5 Docynarene ", WHH 0000000000, KTIN 000000000, 119015, Mockaas r, Teepckas yn, grow Ns 5 Docynarene ", WHH 0000000000, KTIN 000000000, 119015, Mockaas r, Hussel Apfer yn, grow Ns 13 Ne hansenoogenes romaps, pa6or, ycnyr Kone kas 1 Ne hansenoogenes romaps, pa6or, ycnyr Kone kas 1 Wrors: 36 683,00 B rom wachen HUC: 6 950,34 B coro comment: 36 683,00 B rom wachen HUC: 6 950,34 B coro comment: 36 683,00 Decro Hussensonsening 3) wa prywy 30 683/00 pp0 Temperative Docease romants: 36 683,00	Toryware .		-			Inu.re(1000000000000000	
No. Наименования товара, работ, услуг Коли- чество изм. Ед. Цена Сумма 1 Плина СЕЛИМАЦИНО ПНЕ ПОБЛЯЕ О 100 896 600 900 118 21.5 i.0 1.210.00 20.011.00 2 Каклосани (15.5м) 500 500 500.00 3 Плинано, 20120 101 500 500.00 3 Плинано, 20120 101 500 500.00 В том нислов КДС1 6 050,34 6 050,34 5 мася 5 мася Всего в оплати: 36 683,00 5 мася 5 мася 5 мася 5 мася	Toctanuero 000 "Kowneiere", MHH 00 Tocynatere: 000 "Tocynatere", MHH 00 Rocynatere: 000 "Tocynatere: MHH	00000000, KTITI 00000000 000000000, KTITI 0000000	125005,1	Москва г, Тверская "Москва г, Навый	yn, gow Ni S ApGer yn,		
1 Лина СЕЛИМАССИЛО ГИЕ ПОБЛЯЕД (101 89) 600100716 21.3 42 1 210.00 20 013.00 2 Евесоданс (15.54) 100 ort 5.50 500.00 500.00 3 Панима 20100 100 ort 5.50 500.00 500.00 3 Панима 20100 100 ort 5.60 500.00 500.00 4 100 ort 5.60 500.00 500.00 500.00 500.00 8 700 ort 100 ort 56.653.00 100 ort 56.653.00 140 ort 160 ort 1	Ne Наименование товара, 5	абот, услуг К	оли- Ед	L Liena	Сумма		
2 Кактоданся (15 du) 100 dun 1 5.00 500.00 100 dun 2010 Tationare 3 Tationare 3 2 Паними 2010 100 dun 2010 100 dun 2010 36 683,00 1 Tationare 3 1 Harmonia 1 Harmonia <td< td=""><td>1 Towns CERNMAGUARD FINE FIDELIRED (10)</td><td>19-0 400*600*18</td><td>21.3 62</td><td>1 210.00</td><td>20 6 13,00</td><td></td><td></td></td<>	1 Towns CERNMAGUARD FINE FIDELIRED (10)	19-0 400*600*18	21.3 62	1 210.00	20 6 13,00		
1 1100 500000 500000 Итого: 35 653,00 35 653,00 В том инсперида: 6 050,34 4 нательна Всего наименований 3) на сумму 36 653,00 р35. 36 653,00 1 haza Бсего наименований 3) на сумму 36 653,00 р35. 7 арабле 1 haza	2. Exponenter (0.5w)		100 ert.	5.50	560.00	Tational	
Итоля: 36 593,00 В том число НДС: 6 060,34 Всего наименований 3) на сумму 36 653,00 р35. Всего компата: 35 693,00 Всего наименований 3) на сумму 36 653,00 р35. Токаците цисть тысот шестьсот ш	3 Introducto 2020		12 40	300,90	5 300.00	Capital D	NUMBER 2
В том число удс: 6 040,34 Всего коллата: 36 643,00 Бсего наименований 3, на сумму 35 653,00 руб. Толацать шесть тыску шестьорт шестьорт шестьорт сне рубля 60 колекс.				Atera	36 563,00		
Всего коллата: 36 683,00 J. Inco. Бсего наименований 3, на руму 35 663,00 руб. Толацать цесть тыску шестьорт шестьорт шестьорт се рубля 00 колекс.			B	TOM HACTIO HOLC	6 060,34	a 🗘 🖓 🖓 Her	tipoles .
Боего наменования 3, на сумму 35 653/00 руб. Токацать цесть тыски шестьски шестьски шестьски ток рубля 60 колекк.				The state of the s	1000 C 0 0 0 000	The second se	
Tonguate wears taking wears of wears and the provide CO somes.				OCULO & ONTATU:	30 963,00	and the second s	CH :
	ECCO NUMBER OF BUILDING 3, NO CAMPA TO SECT.	0.5051		BORED & DELAMES	36.963,00	4 m	64 4-10

В случае, если роботом будут использоваться данные лишь из одной таблицы, вторая таблица может быть удалена. Для этого достаточно выделить ненужную таблицу и нажать на кнопку « в верхней части рабочей области «Генератора шаблонов».



Вкладка «Атрибуты» позволяет определить и обозначить области документа, данные из которых будут распознаваться роботом и импортироваться в документы «Результаты».

Curristen - Cuar	No only	n ging" - l'emparagi Lustinovos					- B ×			
Isla Ofgature										
D 11 Omportun	notoni	i darin 🕨 21 Origationers georgemen - 🕞 Re	presente +1	Sector 1	alle sta	INCASHING THE PARTY OF THE PART	-			
marali, 1075.		· D D Chelpennie @ house @ Arpet	Ayras D Telines	-	antitalinata (D-D	CR Broom (D) OCR reforman	a contract			
Dista	WEI CYLIN		International Pro-	epa Pielas Vinterra as	ganty special at	edenanteriseen, e even sa prij	+ Q: B Parence 8 A			
SAD SAVE: / Maxee				1000011/00	10		Anno Secures			
Name And	Ree supplieres Cit			10000000	9000000000000		Others 0.000			
1991 22	AREA INCOMPANY AND						Here extreme that can 1			
ODD W	linet in our	r					Righer advises Oncean Faloret genue - Deniet sureComments for			
Doctor		000 "Keenaare", Hitel 000000000, KT	1 2018 F2 In accessoes, s	25005, Mis 119075, M	come ry Temponan Income ry Homail	yn, gosi Nr 9 Apfar yn,				
		Vaces, No. 1.0				+				
-		Нанменование товара, работ, услуг	Kor	и- Ед. 180 кам.	Lama	Суния				
1.0	Inizia CER	NAMACILARIS FIRST FIRST AND RECEIPTING		28.3 ed?	1210,00	30415.00	Alexandra			
	J. Represent (1.64)			10 41	6.50	HE IS	Arguettys			
	-			8.79	Hrans HOC:	34 983,90 E 090,34	0.0000			
Incertor-		DEBINE I, HE SYMME DE BEX.03 (196),		84	aro e oninere:	36 563,00	d and T inter			
Tpages	STR. LOT	сть лысяч внетьсот инстызиент три р	TYGAN OU HOUSE				and an and a second sec			

Для каждого атрибута (выделенной области, данные из которой будут импортированы) необходимо определить Тип данных из выпадающего списка. Например, для атрибута «дата документа» выбираем «Date Time».

 Отврыть ноходны 1005. 	A data > 2) Ofguborani gangueon • [Diposeports (+)	Deferrers arpefyr > 4 Bartenters newer consponse in 100 Herrors referre 100 0/2 form 100 0/2 informed	
Brankeringt OV	ага данного сната доначат состасна с ус. случая на здранодутся настанов санов Гостанцина, ранознателя, г	санның состалын то на сипари . Томар рт ри нагысык арттрог	enges, Paeggeeneered of generate polician ensuel) a nycoarron no (being polician polician ensuel) a norma a mantoprati	+ RJ B Raman S S > Argudyr Sea wearen Argudyr Sea wearen Argudy
SAD "EAHK", r.S	Моская	Gate	000000000	Odnacro 250,210,210,46
Same mony-syname		Cv. Ne	000000000000000000000000000000000000000	Tet
MHH 00000000 OOD "Kownawe Panyseres	xe xmn cooccoccc *	Ox.Ne	000000000000000000000000000000000000000	Propert separate fait One cave 34 Factore Janese Roal Date Tree
Счет на с	оплату № 10 от 01 февра	ля 2018 г.		Aspedyr Chan

При необходимости любой атрибут может быть удален или добавлен новый.

После завершения работы во вкладке «Атрибуты» производится проверка шаблона и выполняется полное сканирование. Нажимаем «Проверить», затем «Выполнить полное сканирование».

Открывается окно «Результат полного сканирования», в котором отображаются распознанные данные из документа в виде заполненной таблицы. 97



Если при проверке обнаружены неточности, то их в любой момент можно исправить.

Chergan - Chermanning grig - Temperap sadimine						18
Bailn Ofpetiensa						
🕄 🖿 1) Ongara eccapad dain 🕨 2) Olgaferara gregorer - 🖬 Ap	овернов 🕂 🕂 Добланов корибут 🛛 🖨 Вылального по	нок ссанорование	8			
Macuatale 100% - D Deedgemanner @ Reoge @ Arped	yrae @ Tatinasae @ Hapresou ratinasa @ OCR Grosse @	D OCR HIGHNAL	Amelyn			
			+ 9	🖞 Openson 🛢 👂		
Бекманией Оклага данного систа саначиет солодно с услови гропивнов случая на геранлоруется надожна гозора на с Госпанција, санотниказори, груп е	 Релультат полного сканародання Релики работы: Простой ракава • Q 0 	Geosette T Doe	eps acts oct	бат. Без инняж	-	0
SAD "EAHI", r Moosee	Applica	Tationus 0 .				
Ease hopy-strates	the Superent	ЗАО ЪАНК", г. москва	les former			
1994 000000000 W7IN 00000000	End answer 10		[avered	Bet Jeneral J		
000 "Kownaww"	East and a 1 02 2018 0 00 00	a feilleaferte	C4, 78			
Donyverens .	Bes weeks "Kovitarius", ihtei 000000000, KTIT 00000000, 1	ARM BOHOHOOG KITS	C+_105	000000000000000000000000000000000000000		
	Ses www. 000 Toxynereis," ИНН 000000000. КПП 0000000	000 videmaneté*				
Счет на оплату № 10 от 01 февраля						
Tioctabuges: ODO "Kownawes", UHH 0000000000, KT						
(Tosynatiania) DOD "Tosynatiania"; KHH 000000000; I						
Na Наименование товара, работ, услуг						

При выборе типа данных «дата документа» дополнительно можно указать «Формат значения» вручную. После чего провести повторную проверку и убедиться в корректности значений.

And Caller				
bein Otpatienta				
👛 1) Отерыть исходный файл	2) Обработать документ • [Проверить +	 Добежить атрибут	gauge .
acurati 100% - 🖉	D Grofipemarter @ Reupe @	Arputtynu 🛞 Talina	nae @ Hepress ratines @ OCR times @ OCR ratiness	Arguelynu
				+ C # Operator 2 8
And a state of the state of the state of the		CONTRACTOR NAME	The second s	Э Атрибат Бел нентон
Enumanue) Crystella come	w Argurdyn Een senares			
Contraction of Contract	Normaniana, passioners	CAR HILL DE HAR ZOUERDEN	INCOME TO COLUMN TY I STATUTE TO STATUTE STATUTE TO STATUTE STATUTE STATUTE STATUTE STATUTE TO STATUTE STATUTE	ibia Serviceo
	and the second s	to the fillen	and and the second s	Repose revealers Texaster (2)
		and the second s	20000000	USANCE 22121021040
SAO "EAHE", r Mocume		Duvin,	0000000	Marganearms or 0
SAO "EAH6", r Mocuma Lass nanysmene		Cx. No	000000000000000000000000000000000000000	Mari sonaracitte or 0 Tati sarean Date Time
SAO "EAH#", r Mockes Less renyversite #9-94.0000000000	x711 00000000	Cv. Na Cv. Na	00000000	Main someworks ox 0 Ten 38-tean Date Time Stocket and second at MV yess
SAO "EArrie", r Mocuse Islee rumy-errere #944 0000000000 OCO "Kowneyver"	KTIT 000000000	Cx. Na Ox. Na	00000000000000000000000000000000000000	Nervanuerter or O Tergenist service Dar De Gruppels and Alfan (1995) Orocane Parlow Danie Danie services and the
SAO "EAnite", r Mocana Lase nanywrana 1994 000000000 OOO "Kostrawer" Potywrana	HULL 00000000	Cv. Na Cv. Na		Men xaturatoritar por 0 Ten galensian Data Tina Oraciant Annual Alf Million Oraciante galense di Alf Million Professe galense di Alfred Professione e ropo Artigone galense di Alfred Professione e ropo Artigone galense di Alfred Professione e ropo

При распознавании данных двух и более таблиц обрабатываемого документа, в окне «Результат полного сканирования» будут отображаться значения всех таблиц. Для просмотра значений по всем таблицам предусмотрена возможность переключения таблиц в выпадающем списке.

Pezyne	тат полного сканирования					-	×
Perces pati	еты Простой рекома - О	Ofecasts.	The	ирх всех окон			
Arpetyne		Telona	0	10			
ibra	3-3-8-44	30034	12	19			-
Sec.man	10	C. HOOKE	1		Second 3		
Ees inner	01.02.2018	a nomen	11	C4_RB			
Ees week	"Kinenawa", idHH 5550000000, KD7 000000000,		000 X/M	Cx_10	0000000000000000000000		
Ex1000	000 Torynemen, MHH 000000000, K111 0000000	000 + 109	runi:				

