

Content AI: автоматизация обработки документов и поиска информации для крупного бизнеса

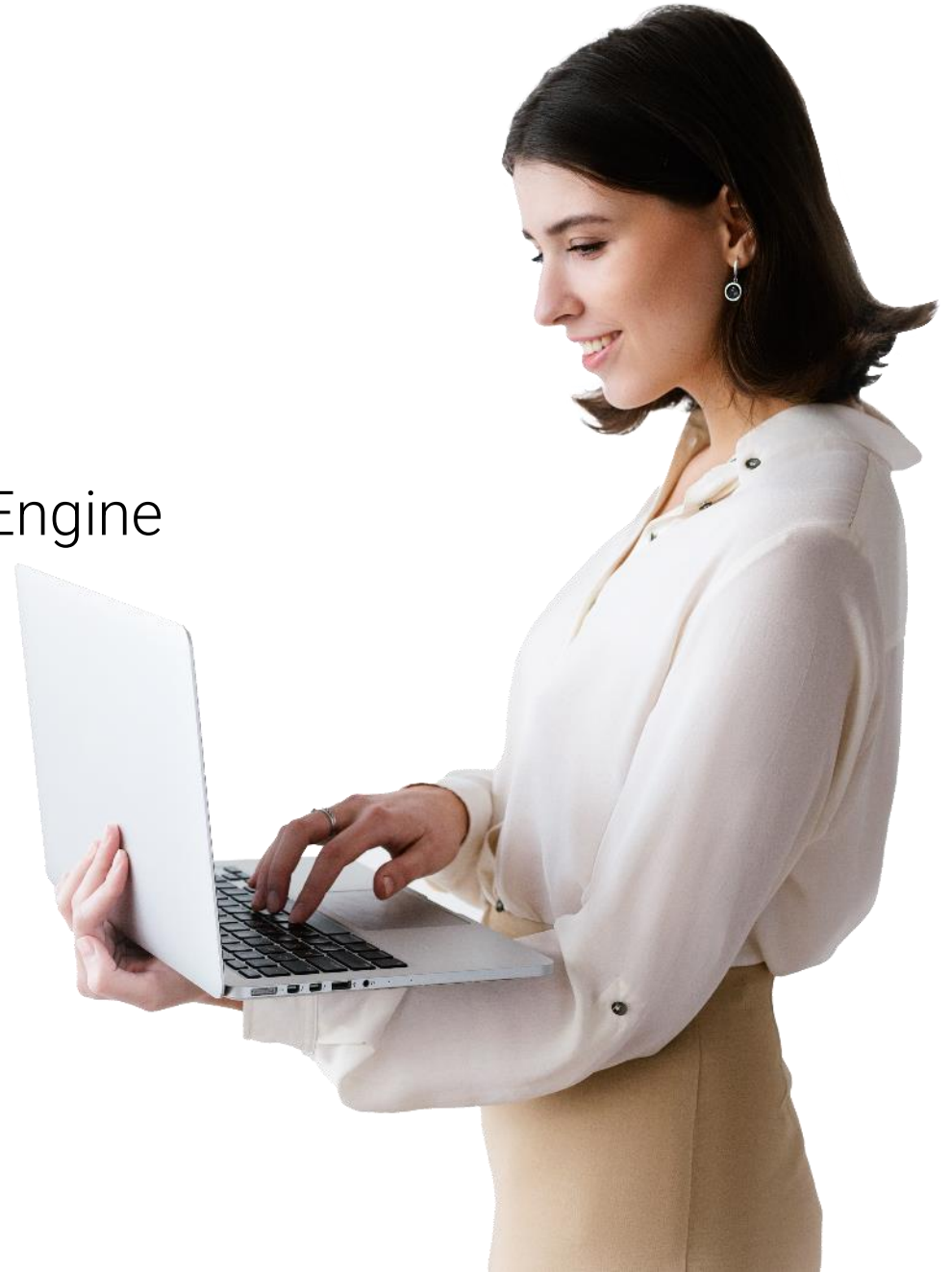
*Иван Козырев,
Менеджер по продажам отдела корпоративных продаж Content AI
Екатерина Ножнёва,
Руководитель проектов отдела интеграции и сопровождения Content AI*

Содержание

➤ Обзор продуктов:

- ContentCapture
- ContentReader Server и ContentReader Engine
- Content AI Intelligent Search

➤ Ваши вопросы



О компании

Content AI (ООО «Контент ИИ») – российская компания, разработчик решений в области интеллектуальной обработки информации.

Content AI выпускает продукты на основе всемирно признанных технологий распознавания текста, классификации документов и обработки естественного языка для крупного бизнеса, государственных организаций и частных клиентов.



100+

сотрудников в Москве и Новосибирске



200+

партнеров на территории России и СНГ



Компания Content AI имеет аккредитацию **Минцифры**



Продукты включены в **Единый реестр российского ПО**



С апреля 2022 года работу с клиентами и партнерами ABBYY в России продолжает компания **Content AI**





CONTENT AI

ABBYY

RU



Ключевые продукты и решения

Массовое ПО

-  ContentReader PDF
-  Lingvo by Content AI






ПО для разработчиков

-  ContentReader Engine
-  PassportReader SDK



Корпоративное ПО

-  ContentCapture
-  ContentReader Server
-  Content AI Intelligent Search



Традиционный документооборот

 ТОРГ-12

Счет-фактуры



Финансовая
и бухгалтерская документация



Приказы



Формы, анкеты, договоры



Отчеты



СНИЛС

2НДФЛ



Кадровые документы



Паспорт



Письма



Входящая корреспонденция



Запросы от гос.
ведомств



Технологии на базе ИИ в продуктах



ML

OCR/ICR

NLP



ContentCapture



Кросс-платформенное решение для автоматизации обработки информации



ContentCapture – универсальная платформа для интеллектуальной обработки информации из любых типов документов: отсканированных бумаг, фотографий, электронных документов, текстов писем и вложений.

Решение распознает, классифицирует документы, извлекает данные, проверяет их корректность и передает в корпоративные информационные системы.



Входит в реестр отечественного ПО и замещает иностранное ПО:

- АБВУУ FlexiCapture
- KofaxCapture
- IBM Datacap

ContentCapture

ОБЗОР ПРОДУКТА



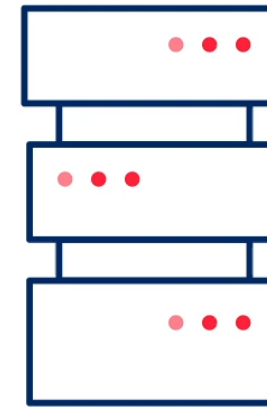
Распознавание

Изображения и документы,
распознаваемые на 200 языках



Понимание

Классификация и извлечение
данных из структурированных и
неструктурированных
документов



Данные

Значимая информация для
принятия бизнес-решений

ContentCapture

Ключевая функциональность

- Единый конвейер для обработки всех видов документов для всех бизнес-процессов
- Гибкая настройка процессов обработки
- Готовые модели обработки для типовых документов
- Простое добавление новых видов документов в существующие процессы
- Масштабирование проектов, поддержка SLA
 - Поддержка распознавания более 200 языков

Что нового?

- Распознавание русского рукописного текста
- Распознавание подписей и печатей
- Добавление новых типов документов через разметку с помощью AI-алгоритмов в дополнение к гибким описаниям

ContentCapture поддерживает Linux в распределённом режиме развёртывания

- **Одинаковая функциональность на разных ОС.** Благодаря переходу на кросс-платформенный технологический стек продукт предоставляет одинаковую функциональность, вне зависимости от используемой ОС.
- **Поддержка гибридной инсталляции.** В процессе импортозамещения на первом этапе на Linux мигрирует только серверная часть, а рабочие станции остаются на Windows.
- **Docker.** Серверная часть решения упакована в docker-контейнеры и готова к работе как на отдельных Linux-машинах, так и в составе docker-кластера.



Распространенные сценарии

ОБРАБОТКА ВХОДЯЩЕЙ ДОКУМЕНТАЦИИ

- оперативная обработка входящих документов, их классификация
- обработка запросов от государственных ведомств (ФОИВ)

ВВОД ДАННЫХ В БУХГАЛТЕРСКИЕ СИСТЕМЫ

- автоматическое извлечение данных из счетов-фактур, накладных, платежных документов и пр. для экспорта в учетную систему

ОТКРЫТИЕ СЧЕТОВ/ КРЕДИТНЫЙ КОНВЕЙЕР

- обработка пакетов документов от физических или юридических лиц, классификация, проверка комплектности, извлечение данных из документов, включая документы удостоверяющие личность и уставные документы организации

ОБРАБОТКА КАДРОВОЙ ДОКУМЕНТАЦИИ

- распознавание и извлечение данных из документов при приеме на работу: паспорт, свидетельство о рождении, водительские права, СНИЛС, 2 НДФЛ и проч.

РАБОТА С ДОГОВОРАМИ

- обработка неструктурированных док-в (договоры, приказы) и судебной документации (решения суда, исполнительный лист)

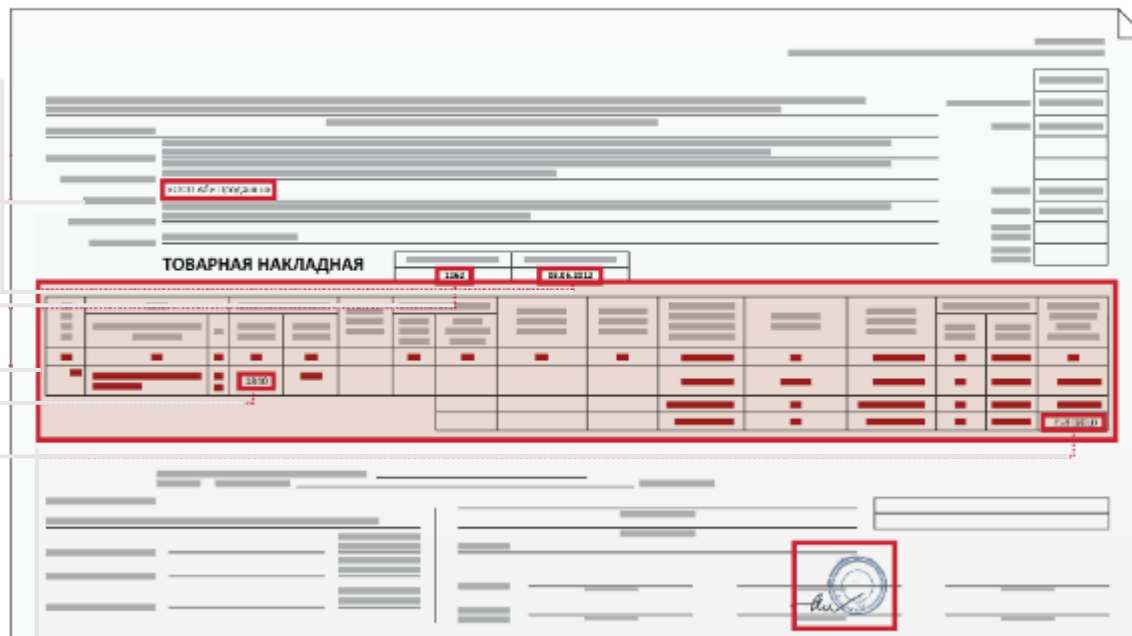
ИМПОРТОЗАМЕЩЕНИЕ

- миграция с аналогичных зарубежных решений, ушедших с российского рынка:
- ABBYY FlexiCapture, Kofax Capture, IBM DataCap и прочих.

Демонстрация работы извлечения значимой информации из документов

Извлечение данных

- Номер
- Дата
- Поставщик
- Реквизиты
- Количество
- Сумма
- ...
- Извлечение табличной части документа



Проверка данных

- Форматные проверки
- Нормализация наименования номенклатурной позиции
- Проверка наличия печати и подписи
- Контроль сумм и НДС
- Проверка по справочникам организаций

Автоматизация ввода данных в бухгалтерии

ЦЕЛЬ: сокращение издержек на обработку документации

Задачи

- Автоматизация извлечения необходимых данных из документов, внутри документные проверки – наличие подписей, корректность расчетов
- Оптимизация процесса работы с входящими бумажными документами в бухгалтерии

Результаты

Экономия

Сокращение издержек на обработку первичной документации в 2 раза

Лояльность

Повышение уровня лояльности партнеров благодаря росту скорости обслуживания



Снижение рисков

Снижение рисков предъявления претензий и наложения штрафных санкций со стороны налоговых органов



ContentReader Server ContentReader Engine



Решения для распознавания,
конвертации больших объемов
документов и сравнения версий
договоров

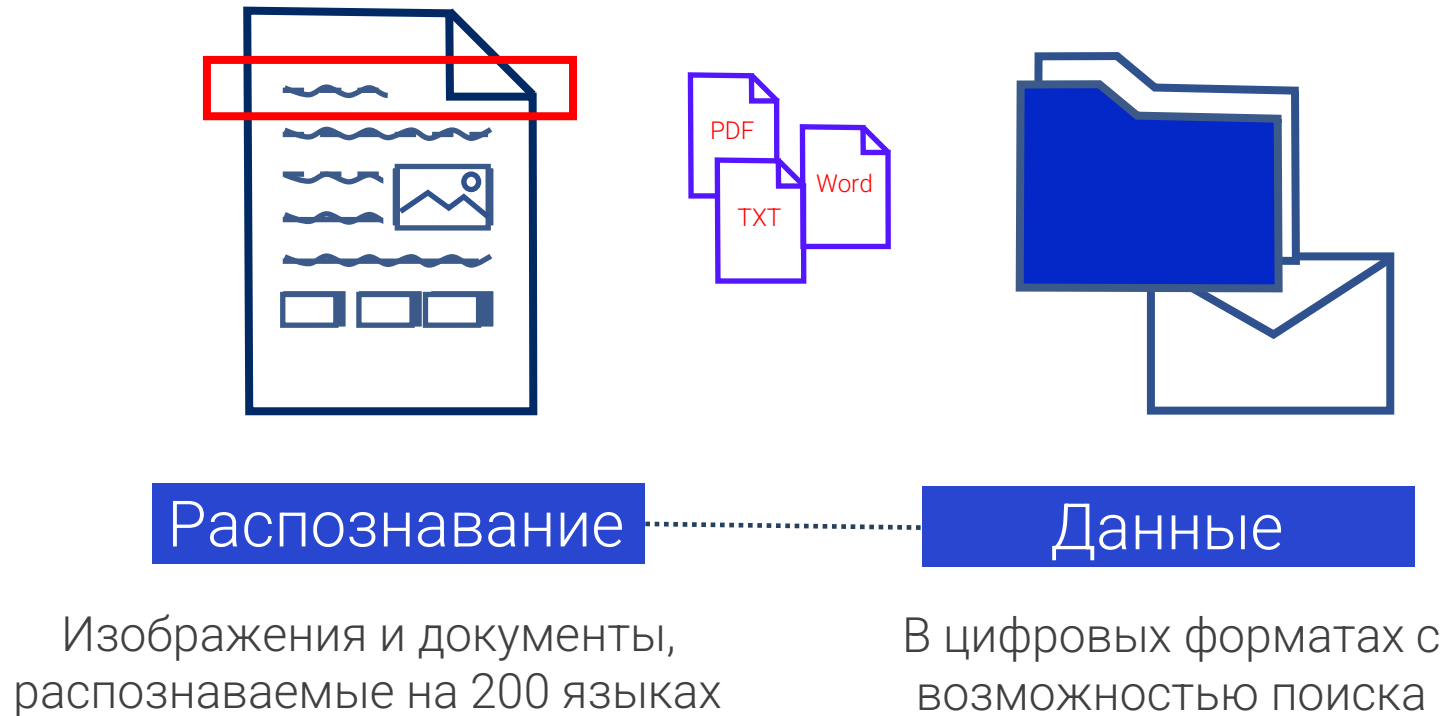
Корпоративные продукты для распознавания и преобразования файлов в PDF и другие электронные редактируемые форматы. Технологии помогают пользователям быстро отслеживать изменения, проверять целостность документов и предотвращать мошенничество.



Входит в реестр
отечественного
ПО и замещает
иностранное ПО:

- ABBYY
Recognition
Server/FineReader
Server
- IRISXtract

ContentReader Server



Распространенные сценарии

ЭЛЕКТРОННЫЙ АРХИВ

- преобразование больших объемов документов в легкодоступные цифровые форматы

ЦЕНТРАЛИЗОВАННЫЙ СЕРВИС КОНВЕРТАЦИИ

- автоматизированная конвертация всех документов из отсканированных документов в редактируемые форматы для сотрудников организации

СОЗДАНИЕ ЦИФРОВЫХ БИБЛИОТЕК

- преобразование исторических документов в доступные для поиска файлы PDF, XML или EPUB

ВВОД ВХОДЯЩИХ ДОКУМЕНТОВ В РЕЖИМЕ «КАНЦЕЛЯРИЯ»

- сканирует документы и регистрирует их в системе, выделяя необходимые атрибуты документа (номер входящего документа, дата и т. д.)

СРАВНЕНИЕ ДВУХ ВЕРСИЙ ДОКУМЕНТОВ

- интегрируется в корпоративное ПО. Автоматически отслеживает изменения, проверяет целостность документов и помогает предотвращать мошенничество.

ИМПОРТОЗАМЕЩЕНИЕ

- миграция с аналогичных зарубежных решений, ушедших с российского рынка: ABBYY FineReader/Recognition Server, IRISXtract и прочих.

Демонстрация: сравнение документов

и приложениям к нему, а также осуществлять какую-либо аудио/видеозапись без согласия Исполнителя.

3. ЦЕНА ДОГОВОРА И ПОРЯДОК РАСЧЕТОВ

3.1. Стоимость услуг Исполнителя и порядок оплаты указываются в Приложении №1 к Договору.

3.2. Если иное не указано в Приложении №1, оплата производится Заказчиком путем перечисления денежных средств на расчетный счет Исполнителя в размере 100 % от стоимости услуг в течение **14 (четырнадцать)** календарных дней с даты выставления счета.

3.3. Датой оплаты считается дата поступления денежных средств на корреспондентский счет Банка Исполнителя.

3.4. В случае неоплаты услуг в срок, установленный в Договоре, Исполнитель вправе, уведомив Заказчика, отказаться от исполнения обязательств по Договору без выплат/компенсаций Заказчику каких-либо расходов/убытков.

4. ПОРЯДОК СДАЧИ И ПРИЕМКИ УСЛУГ

4.1. Исполнитель в течение 5 (пяти) рабочих дней с момента окончания оказания услуг, направляет Заказчику подписанный со своей стороны Универсальный передаточный документ (далее - УПД) в 2 (двух) экземплярах, а так же отчет об оказанных услугах, который подтверждает оказание услуг перечисленных в Приложении №1 к Договору, включающий в себя в том числе ссылки на сайты, фото,.

4.2. Заказчик обязуется в течение 5 (пяти) рабочих дней со дня получения УПД направить подписанный экземпляр Исполнителю, либо в тот же срок направить в адрес Исполнителя письменный мотивированный отказ от его подписания.

4.3. В случае получения мотивированного отказа Заказчика от подписания УПД по настоящему Договору, Стороны согласовывают дальнейшие действия по исполнению своих обязательств в рамках настоящего Договора.

4.4. В случае уклонения или немотивированного отказа Заказчика от подписания УПД в течение 5 (пяти) рабочих дней со дня его получения, УПД считается подписанным, а услуги оказанными в полном объеме, надлежащим качеством и принятыми Заказчиком без претензий.

5. СРОК ДЕЙСТВИЯ ДОГОВОРА И ПОРЯДОК РАСТОРЖЕНИЯ

5.1. Настоящий Договор вступает в силу с момента подписания Сторонами на бумажном носителе и действует до полного и надлежащего выполнения Сторонами своих обязательств по настоящему Договору. В случае подписания Договора посредством системы ЭДО (электронного документооборота) с использованием УКЭП (усиленных квалифицированных электронных подписей) Стороны руководствуются положениями пп. 9.9 – 9.11 Договора.

5.2. Настоящий Договор может быть досрочно расторгнут:

5.2.1. По соглашению Сторон путем подписания соглашения о расторжении;

5.2.2. По инициативе одной из Сторон с обязательным письменным уведомлением другой Стороны.



и приложениям к нему, а также осуществлять какую-либо аудио/видеозапись без согласия Исполнителя.

3. ЦЕНА ДОГОВОРА И ПОРЯДОК РАСЧЕТОВ

3.1. Стоимость услуг Исполнителя и порядок оплаты указываются в Приложении №1 к Договору.

3.2. Если иное не указано в Приложении №1, оплата производится Заказчиком путем перечисления денежных средств на расчетный счет Исполнителя в размере 100 % от стоимости услуг в течение **60 (шестидесяти)** календарных дней с даты выставления счета.

3.3. Датой оплаты считается дата поступления денежных средств на корреспондентский счет Банка Исполнителя.

3.4. В случае неоплаты услуг в срок, установленный в Договоре, Исполнитель вправе, уведомив Заказчика, отказаться от исполнения обязательств по Договору без

3.2. Если иное не указано в Приложении №1, оплата производится Заказчиком путем перечисления денежных средств на расчетный счет Исполнителя в размере 100 % от стоимости услуг в течение **60 (шестидесяти)** календарных дней с даты выставления счета.

оказание услуг перечисленных в Приложении №1 к Договору, включающий в себя в том числе ссылки на сайты, фото,.

4.2. Заказчик обязуется в течение 5 (пяти) рабочих дней со дня получения УПД направить подписанный экземпляр Исполнителю, либо в тот же срок направить в адрес Исполнителя письменный мотивированный отказ от его подписания.

4.3. В случае получения мотивированного отказа Заказчика от подписания УПД по настоящему Договору, Стороны согласовывают дальнейшие действия по исполнению своих обязательств в рамках настоящего Договора.

4.4. В случае уклонения или немотивированного отказа Заказчика от подписания УПД в течение 5 (пяти) рабочих дней со дня его получения, УПД считается подписанным, а услуги оказанными в полном объеме, надлежащим качеством и принятыми Заказчиком без претензий.

5. СРОК ДЕЙСТВИЯ ДОГОВОРА И ПОРЯДОК РАСТОРЖЕНИЯ

5.1. Настоящий Договор вступает в силу с момента подписания Сторонами на бумажном носителе и действует до полного и надлежащего выполнения Сторонами своих обязательств по настоящему Договору. В случае подписания Договора посредством системы ЭДО (электронного документооборота) с использованием УКЭП (усиленных квалифицированных электронных подписей) Стороны руководствуются положениями пп. 9.9 – 9.11 Договора.

5.2. Настоящий Договор может быть досрочно расторгнут:

5.2.1. По соглашению Сторон путем подписания соглашения о расторжении;

5.2.2. По инициативе одной из Сторон с обязательным письменным уведомлением другой Стороны.

Оптимизация процесса сравнения договоров

ЦЕЛЬ: Повышение эффективности обработки договоров в организации

Задачи

- Оперативно обнаруживать различия в содержании между двумя версиями документов, включая отсканированные версии договоров
- Сократить время сотрудников, которое уходит на этот процесс в ручном режиме

Результаты

Сокращение времени

Ускорился процесс согласования договоров с партнерами и клиентами

Вовлеченность

Специалисты, ответственные за работу с договорами, могут посвятить больше времени решению более важных задач



Снижение рисков

Решение позволило уменьшить количество ошибок в документах и избежать проблем, связанных с недобросовестными контрагентами

Content AI Intelligent Search



Корпоративная поисковая система

- Сквозной поиск по всем разрозненным корпоративным системам предприятия в едином окне
- Поиск по смыслу и точному совпадению
- Встроенная технология OCR для поиска по графическим файлам
- Аналитика данных в источниках
- Продукт включен в реестр отечественного ПО

Источники:

- Файловый сервер
- SharePoint
- Confluence
- Jira
- СУБД
- СЭД
- Битрикс24
- Каталог пользователей Active Directory
- Microsoft Exchange
- Azure DevOps Server (TFS)
- и др.



Распространенные сценарии

СКВОЗНОЙ ПОИСК ПО ВСЕМ
ИСТОЧНИКАМ ДАННЫХ

- упрощает и ускоряет поиск полезной информации в документах прошлых лет

ПОИСКОВЫЙ ПОРТАЛ
ДЛЯ УСКОРЕНИЯ ЗАПУСКА
НОВЫХ ПРОДУКТОВ

- поиск по нереализованным ранее решениям, патентам конкурентов, неудачным опытам, по истории взаимодействия с партнерами в различных проектах и проч., чтобы избежать дублирования ошибок

ПОИСК ДЛЯ ОТДЕЛА
ИНФОРМАЦИОННОЙ
БЕЗОПАСНОСТИ

- корпоративный поиск по материалам в различных форматах, включая почтовые архивы, для проведения внутренних финансовых расследований

ПОИСКОВЫЙ ПОРТАЛ ДЛЯ
ЮРИСТОВ

- в кратчайшие сроки собрать актуальную и точную информацию по текущему делу, а также о похожих кейсах прошлых лет

ПОИСКОВЫЙ ПОРТАЛ ДЛЯ
БУХГАЛТЕРИИ

- увеличивает скорость поиска и обработки бухгалтерских и финансовых документов и подготовки отчетов

ПОИСК ЭКСПЕРТОВ ВНУТРИ
КОМПАНИИ

- сохранение знаний и компетенций внутри компании
- поиск узких специалистов по заданным критериям

Демонстрация: поиск

The screenshot displays the Content AI Intelligent Search interface. At the top, the search bar contains the query "шаблон презентации" and the "Найти" (Find) button. The interface is divided into three main sections: filters on the left, search results in the center, and a detailed view of the selected document on the right.

Filters (Фильтры):

- Источник (Source):** Портал MS Sharepoint (103), Файловый сервер (69), Портал Confluence (29).
- Библиотека / Пространство (Library / Space):** Marketing (103), Emerging Markets (3A) (6), DataCapture (5), Creative Group (5), Knowledge Base (5).
- Тип элемента (Element Type):** Документ из библиотеки (103), Веб-страница (16), Вложение (12), Запись в блоге (1).
- Должность автора (Author Role):** Бизнес-аналитик (50), Руководитель департамента (5), ТАМ (4), Разработчик (4), Руководитель проектов (3).
- Автор изменений (Change Author):** 35.

Search Results (Найдено: 201):

- Шаблон Presentation_16x9 2022.pptx** (Selected): ...шаблона 13.04.2022 Представляем Вам новый шаблон презентации * | Тема презентации | XX/XX/XX Примеры...
22.04.2022, 15:32:16 | \\fileserv\Marketing Department\Product Management Group\Publishing Service\Deposit\Temp\FOR_NEW_BRANDBOOK\example\Шаблон Presentation_16x9 2022.ppt
- Portfolio_Management_Presentation_Template.pptx**: ... How to work with this presentation template Please, edit only slides from the sections "START" and...
- Portfolio_Management_Presentation_Template.pptx**: ... How to work with this presentation template Please edit only slides from the sections "START" and "EDITABLE"...
- Presentation Template and Guide**: ...for creating a presentation are set out in a specialized guide, Corporate Template Brand Guide...
18.03.2021, 15:52:23 | https://wiki.contentai.ru/display/DSGN/Presentation+Template+and+Guide
- 10027r_Presentation_Template.pptx**: 13.03.2019, 14:21:29 | \\fileserv\Marketing Department\Product Management Group\Publishing Service\Deposit\10027r\presentation\work\10027r_Presentation_Template.pptx
- Presentation_template_2021.pptx**: 24.06.2021, 08:43:43 | \\fileserv\Marketing Department\Product Management Group\Publishing Service\Deposit\Presentation\Old template\Presentation_template_2021.pptx
- Portfolio_Management_Presentation_Template.pptx**: ...DISCLOSURE The information contained within this presentation is for internal use only and should not...
14.04.2020, 18:53:38 | \\fileserv\Marketing Department\Product Management Group\Publishing Service\Deposit\11671\w

Document Preview (Шаблон Presentation_16x9 2022.ppt):

- Общая информация (General Information):** Изменен: 22.04.2022, 15:32:16. Ссылка / путь: \\fileserv\Marketing Department\Product Management Group\Publishing Service\Deposit\Temp\FOR_NEW_BRANDBOOK\example\Шаблон Presentation_16x9 2022.ppt
- Текст документа (Document Text):** * | Тема презентации | XX/XX/XX
Образец заголовка
Презентация по правилам использования корпоративного шаблона
13.04.2022
Представляем Вам новый шаблон презентации
* | Тема презентации | XX/XX/XX
Примеры и правила верстки презентаций
Примеры шаблонов
Содержание презентации
* | Тема презентации | XX/XX/XX
Примеры и правила верстки презентаций
* | Тема презентации | XX/XX/XX

At the bottom, there is a pagination bar showing "1 2 3 4 5 6 7 8 9 ... 29" and a footer with "Язык запросов", "Панель предпросмотра", and the "CONTENT AI" logo.

Поисковый портал для запуска новых продуктов

ЦЕЛЬ: сократить время на создание продукта

Задачи

- Поиск информации по нереализованным ранее решениям, патентам конкурентов, неудачным опытам, чтобы избежать дублирования ошибок
- Ускорить поиск по истории взаимодействия с партнерами в различных проектах

Результаты

400 000

документов, включая отчеты о научно-технической деятельности, патенты, диссертации, и т.п. доступны через сквозной поиск

Процесс разработки продуктов

был оптимизирован благодаря реализации поиска по ключевым источникам информации в организации: файловому серверу, MS Sharepoint, Confluence



В 3 раза

сократилось время подготовки заявки на новую продукцию

Автоматизация документооборота с Content AI



ContentCapture

- Обработка бухгалтерских и финансовых документов
- Ввод кадровой документации
- Анализ резюме соискателей, договоров и других документов
- Классификация входящего потока обращений



ContentReader Server

- Электронный архив
- Централизованный сервис конвертации документов для сотрудников
- Поисковый портал по всем источникам организации
- Аналитика по данным поиска



Спасибо за внимание!

Обратиться за триальной версией
продуктов:

dmitriy.ponomarev@softline.com

+7 (495) 232-00-23