



# FormXtra<sup>®</sup> Capture

Комплексное решение для распознавания документов

## Готовое интеллектуальное решение по распознаванию документов

*FormXtra Capture* — это решение по обработке документов и их распознаванию, разработанное для предприятий, которые нуждаются в продукте, обладающем высокой производительностью, адаптационной способностью и простотой в эксплуатации. Решение предназначено для обработки больших объемов сложных документов, имеет встроенный распознаватель банковских чеков и позволяет пользователю индивидуально настроить правила обработки результатов.

Продукт имеет удобный интерфейс и встроенную систему интерактивного тестирования распознавания с вычислением процента ошибок для отдельных полей и форм целиком. С помощью этого продукта компании из самых разнообразных областей деятельности (страховые, финансовые, торговые компании, банки) могут осуществить автоматизацию полного цикла обработки данных, получив возможность контролировать процессы распознавания и ввода данных. При этом *FormXtra Capture* способствует поддержанию уровня безопасности, предоставляя развитый инструмент для разграничения прав доступа. Масштабируемая модульная архитектура и обширная функциональность позволяют с легкостью встроить *FormXtra Capture* в уже работающую систему обработки данных, приспособить продукт к нуждам конкретного вида бизнеса, не тратя существенных дополнительных средств.

В результате пользователь имеет хорошо налаженный инструмент для автоматической обработки документов, обеспечивающий высокую точность распознавания и скорость обработки. *FormXtra Capture* предоставляет все функции обработки документов: ввод, классификацию документов, их распознавание, валидацию данных, вывод в заданной форме. Такая функциональность является стандартом в современных системах интеллектуального распознавания документов.

## Динамическая классификация документов

### **Расширенная функциональность в обработке документов любого типа**

*FormXtra Capture* идентифицирует тип документа, используя при классификации усовершенствованные, гибкие критерии, а не только жестко фиксированные шаблоны, предопределяющие размеры и вид того или иного элемента документа (таблицы, «шапки» или другие специфические части документа).

Использование сложных алгоритмов классификации позволяет точно отнести к одной категории документы, содержащие одинаковый набор полей, даже если эти поля расположены в разных местах. Для идентификации поля может использоваться распознавание заголовка поля.

## Распознавание любых типов данных

### **OCR, ICR и Parascript ICR (для распознавания рукописного текста, написанного в свободной, естественной манере)**

Благодаря использованию технологий оптического посимвольного распознавания (OCR) и интеллектуального распознавания (ICR), *FormXtra Capture* читает практически все типы данных: печатный текст, рукопечатный текст (как заключенный, так и не заключенный в рамки), слитный рукописный текст (написанный в свободной, естественной манере), 1D/2D штрихкоды, метки (OMR).

Смешанный  
сбор данных  
со структурированных и  
полуструктурированных  
документов

Широкий спектр  
возможностей  
по обнаружению  
и распознаванию полей  
разных типов

Усовершенствованная  
технология  
валидации данных

Встроенное  
обнаружение адреса  
и распознавание

#### Повышает гибкость рабочего процесса

---

Программа может находить и обрабатывать как фиксированные поля (местоположение которых на форме фиксировано), так и динамические поля (местоположение которых заранее не известно; известно лишь, что поле должно присутствовать на одной из страниц документа определенного типа).

Такая возможность увеличивает гибкость обработки больших объемов документации и сокращает количество классов документов, которые необходимо описать и обрабатывать отдельно.

#### Обнаружение и обработка полей любых типов в любом месте документа.

---

Обрабатывает структурированные и полуструктурированные документы, а также формы, которые не соответствуют предварительно заданному шаблону, например, уменьшенные в размере факсы и таблицы. *FormXtra Capture* позволяет описывать сложные схемы взаимного расположения разных полей в документе, поля сложной внутренней структуры.

Уникальной особенностью является возможность обнаружения и последующего распознавания рукописных полей на документах с жесткой структурой — например, на счетах и накладных, выписанных вручную.

#### Механизмы для обеспечения высокой защищенности обработки конфиденциальной информации (финансовой, медицинской и прочей)

---

*FormXtra Capture* предоставляет развитый инструмент для описания сложных схем валидации и кроссвалидации данных – от задания простейших правил (проверка попадания значения поля в заданный диапазон), до сложных сценариев, предписывающих двойную слепую обработку поля с аудитом расхождений и исправлением ошибок. При этом, в целях соблюдения конфиденциальности (например, при обработке персональных данных) есть возможность показывать аудитору не всю страницу (или документ), а только лишь участок документа, соответствующий одному полю.

#### Обработка входящей почты.

---

*FormXtra Capture* позволяет обрабатывать входящую почту и ее содержимое, что позволяет находить адреса на конверте, либо в теле документа, чтобы использовать эти данные для переадресации.

В комбинации с распознаванием полностраничных данных важная корреспонденция может быстро перенаправляться соответствующему получателю; адресные данные могут сопоставляться с другими документами, таким как заказы на поставку, счета, платежные поручения, денежные переводы.

## Встроенное распознавание банковских чеков

## Широкий спектр возможностей по вводу и выводу данных

## Удобный интерфейс для настройки системы обработки данных и ее администрирования в процессе эксплуатации

### Обработка записей о платежах и денежных переводах.

*FormXtra Capture* позволяет обрабатывать банковские чеки в пакете других документов с поддержкой приложений, которые обрабатывают, сопоставляют и анализируют записи о платежах и денежных переводах. Решение позволяет воспользоваться высококачественными механизмами распознавания чеков, созданными компанией *Parascript* для специфических банковских приложений, и использовать их для повышения точности обработки форм.

Новые функциональные возможности и расширенный API обеспечивают широкие возможности для интеграции с другими системами. *FormXtra Capture* предоставляет вывод информации в SharePoint, а также в любые базы данных в виде общепринятых форматов данных.

### Графический пользовательский интерфейс для представления результатов и тестирования в реальном времени.

*FormXtra Capture* снабжен удобным графическим пользовательским интерфейсом для настройки и тестирования системы по распознаванию документов, а также ее администрированию в процессе эксплуатации. Это позволяет создать жизнеспособное решение автоматической обработки данных без дополнительных затрат на программирование и интеграцию нескольких продуктов по обработке и распознаванию документов.

Интерфейс позволяет:

- получить упрощенный анализ в виде графических диаграмм

В режиме реального времени на экран администратора выводится полная статистика о состоянии обрабатываемых документов — от их сканирования и распознавания до перепроверки аудитором и выводом в базу данных.

- получить полный, развернутый анализ

Просмотр результатов идентификации типа документа, выделения полей, результатов распознавания, установки пороговых значений коэффициента уверенности распознавания, сравнение результатов распознавания с эталонными данными, моментальный просмотр результатов распознавания.

- разрабатывать и тестировать скрипты (бизнес-правила) в пределах *FormXtra Capture*

Бизнес правила для создания описаний форм, в том числе через .NET API, встроенные в решение, позволяют еще больше адаптировать процесс обработки документов к нуждам пользователя. При этом не требуются дополнительные внешние приложения (интегрированные среды разработки). Все работы по созданию, тестированию, компиляции выполняются в пределах *FormXtra Capture*.

## Техническая спецификация продукта

### Включенные программные компоненты:

- Form Definition Studio.
- Дополнительные программные компоненты распознавания представлены в виде библиотек .NET API.
- Словари и почтовые базы адресов.

### Системные требования:

- Операционные системы: Windows® 7, Windows® 8, Windows® 2008 Server or higher.
- CPU: Dual-core 3 GHZ or higher
- RAM: 4 GB (8 GB рекомендуется).

Дисковое пространство: 1 GB свободного дискового пространства для установки приложения.

Дополнительное место на диске требуется для хранения файлов изображений, предназначенных для создания и настройки описания документов.

### Поддерживаются следующие форматы входных данных:

- Формат изображений: черно-белые TIFF или BMP
- Разрешение изображений: 200 – 300 dpi

### Поддерживаются следующие форматы выходных данных:

- ASCII текстовый файл
- XML
- Microsoft SharePoint 2010 Standard или Enterprise
- RDBMS через ODBC

## Контактная информация



### IDR

Харьков, Украина

Телефоны:

+38 (097) 559-19-29

+38 (095) 619-18-17

Электронная почта:

[info@idr.in.ua](mailto:info@idr.in.ua)



### Parascript

Лонгмонт, штат Колорадо, США

Телефоны:

+1 (303) 381-31-00

+1 (888) 225-01-69

Электронная почта:

[info@parascript.com](mailto:info@parascript.com)