

Коробочное ПО для ввода документов

ИВАН САВВАТЕЕВ

Компания Cognitive Technologies (www.cognitive.ru) объявила о выпуске коробочного продукта Cognitive Forms Network (CFN), предназначенного для массового ввода стандартных форм документов.

ДОКУМЕНТООБОРОТ

До последнего времени фирма делала упор на проектные решения, однако проведенные Cognitive исследования показали, что сегодня повышенный спрос на подобные системы имеется в сегменте средних предприятий (ввод от 10 до 50 тыс. документов в неделю). В то же время реализация заказного проекта для них не всегда приемлема из-за высоких затрат. Удовлетворить потребности таких заказчиков и призвано новое коробочное решение, с помощью которого Cognitive надеется занять около 60% данного сегмента рынка, оцениваемого ею в 20 млн. долл.

CFN предназначен для массового ввода форм, т. е. документов, содержащих фиксированный набор информационных полей (анкет, квитанций, деклараций и т. п.), с применением технологий сканирования и оптического распознавания. ПО обеспечивает обработку печатного и рукопечатного текста, флажков, маркированных позиций, штрихкодов. CFN способен также разбирать отдельные поля на составляющие их элементы, например анализировать российские адреса с учетом вариантов их написания.

Решение имеет модульную структуру, что позволяет строить технологические линии ввода форм, соответствующие структуре информационных потоков организации и учитывающие специфику самих документов. В состав системы входят модули сканирования, распознавания, верификации, экспорта и администрирования. Технологическая линия может включать десятки рабочих мест, взаимодействующих между собой и ин-

тегрированных с информационными системами предприятия.

CFN поддерживает ввод разных типов форм в одном потоке без предварительной сортировки документов, автоматически распознавая тип каждого из них. Его базовый вариант обеспечивает экспорт введенной информации в форматы DBF, XML, TXT, а также интеграцию с системой "Евфрат-Документооборот".

С "коробкой" поставляется несколько готовых форм, с помощью которых можно протестировать и изучить систему, но совсем избежать настройки не удастся, поскольку практически каждая организация использует свои образцы форм. Для настройки на конкретные виды документов предусмотрен пакет разработчика — Cognitive Forms Developer Kit (CFDK). Он обеспечивает полный цикл настройки технологической линии ввода документов, включая:

- формирование макетов машиночитаемых форм для печати и предварительного заполнения, в том числе в виде интерактивного файла формата PDF;

- разработку шаблонов распознавания, при помощи которых система определяет тип документа и выделяет поля ввода данных;

- подключение механизмов распознавания, соответствующих стилю заполнения полей ввода с учетом алфавита, языка и разного рода синтаксических ограничений;

- разработку и подключение системы контекстного контроля, обеспечивающей автоматизированную проверку правил заполнения документа непосредственно на этапе преобразования его из бумажного вида в электронный. Предусмотрено множество способов автоматического контроля информации, в том числе проверка по словарям и перекрестная проверка значений полей из разных документов.

Для предприятий, планирующих ав-

томатизацию ввода относительно небольшого количества типов форм, разработка этих форм и настройка CFN может быть осуществлена специалистами Cognitive или ее партнеров. Тем компаниям, которые работают со множеством различных типов форм, может оказаться более целесообразным приобрести CFDK (он поставляется отдельно от CFN и стоит порядка 7 тыс. долл.) и обучить своих специалистов самостоятельной настройке CFN.

На сегодняшний день Cognitive предлагает пять вариантов решения: Cognitive Forms Network 2006 Limited Edition (LE), Standard Edition (SE), Profession Edition (PE), Extended Edition (XE) и Extended Edition Plus (XE+). Версия LE (1450 долл.) предназначена для ввода до 5000 стр. в неделю и рассчитана на автоматизацию только одного рабочего места оператора. Остальные версии имеют средства групповой работы; стоимость их начинается с 6500 долл. (версия SE).

Единственным аналогом CFN пока является пакет FormReader Enterprise компании ABBYY (www.abbyy.ru). Однако, по мнению руководителя департамента систем массового ввода Василия Постникова, по ряду характеристик он уступает решению Cognitive, что во многом объясняется "генеалогией" этих продуктов: FormReader "вырос" из настольного решения, в то время как CFN явился результатом "упаковки в коробку" крупных заказных разработок. В частности, CFN реализует полностью распределенную обработку документов (сканирование, распознавание, верификация и экспорт могут осуществляться на различных компьютерах, удаленных друг от друга на любые расстояния и связанных между собой локальной сетью или через Интернет, причем этапы обработки могут быть разнесены во времени) и допускает подключение новых модулей без остановки всего технологического процесса. В ком-

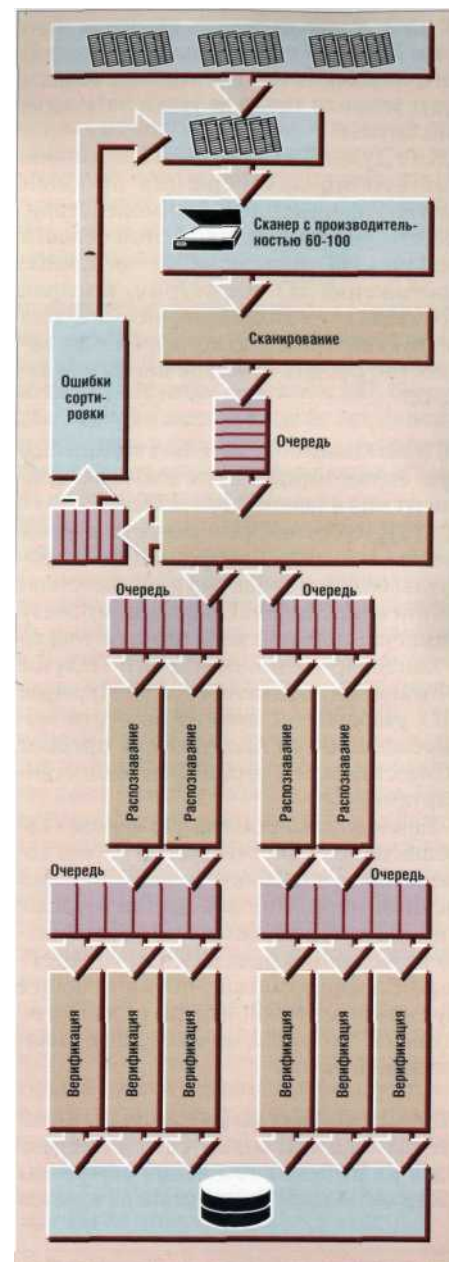


Схема обработки документов в CFN

пании считают, что CFN предпочтут фирмы, нуждающиеся в обработке большого количества форм.