ABBYY®

FlexiCapture™ 1.5

Intelligent Technology for Semi-structured Forms and Documents

Overview: An Answer for Semi-Structured Document Processing

Each day companies around the world struggle with the monumental task of processing millions of paper documents -all of which need to be input into the computer in the quickest, most cost-effective way possible. While today's form processing systems have advanced significantly, the available options for processing semi-structured documents, or "flexible forms," have remained limited.

traditional methods for capturing semi-structured documents have required many hours, days, or months of customization - creation of multiple form templates for each form layout, form redesign, and technology adaptation. Or alternatively, manual input by an operator. ABBYY now offers an alternative.

ABBYY FlexiCapture Studio provides an easy-to-use solution for processing documents that contain similar sets of information, but have varied form layouts and structures. Designed as a companion product for ABBYY FormReader and ABBYY FineReader Engine, ABBYY FlexiCapture Studio enables the creation of intelligent form layouts that work together with ABBYY recognition software to detect key form field data despite variations in layout and placement. It is an ideal solution for processing a wide variety of varied, semi-structured documents in the following areas:

• Finance\Banking (bank payment orders, financial statements, credit notes, loan\mortgage documents)

• Accounting (invoices, purchase orders, payment slips)

• Legal (template-based contracts, lease agreements)

• Insurance (benefit application forms, claim forms)

• Transportation (shipping orders, delivery notes, customs declarations)

ABBYY FlexiCapture: Intelligent Technology

How does a human operator enter data from semi-structured forms into a computer? When looking for data, the operator looks at the entire form as a whole and then analyses various form elements and individual pieces of information such as: field captions, or units of measurement. The operator then looks at information surrounding the form fields to determine which information to enter into which part of the database.

ABBYY FlexiCapture uses artificial intelligence to enable ABBYY recognition technologies (ABBYY FormReader and ABBYY FineReader Engine) to act in a similar fashion. FlexiCapture is based on the principles of Intelligent, Purposeful, and Adaptable (IPA) recognition. IPA imitates the ways in which humans recognize objects. With user-defined descriptions, FlexiCapture guides ABBYY recognition engines toward finding the required fields on forms with varied layouts. As a result, the recognition software can find a key field and piece of data, such as "date," on various different forms - **no matter where it is located on the paper** (at the top left, bottom right, etc.).

How FlexiCapture Studio Works: The FlexiLayout

FlexiCapture Studio extends the data capture and recognition capabilities of ABBYY FormReader and ABBYY FineReader Engine beyond traditional form processing. It achieves this by enabling simple creation, testing, and debugging of a FlexiLayout, a formal description, or "map" of the fields and variable layouts of a certain set of forms.

Customers can work with FlexiLayouts by:
• **Using the FlexiCapture Studio interface** - the intuitive dialogues require no coding or programming experience.
• **Fine-tuning the FlexiLayout** - using FlexiCapture's special template creation programming language.

FlexiLayouts describe the logic of data extraction. Each FlexiLayout consists of user-defined *elements* and *blocks.* *Elements* in the FlexiLayout serve as a set of rules or descriptions of objects and their relationship within the document. They help ABBYY recognition technology find the necessary data fields during form processing and document recognition. *Elements* include static text, character strings, white spaces, etc. *Blocks* describe the elements from which the data must be captured - they are the areas from which ABBYY recognition technology must capture data.

High accuracy and speed is achievable because FlexiCapture uses a unique algorithm to detect data fields by analyzing whole sets of form elements. This approach is dramatically different from that of traditional form processing technology which analyses separate elements one by one. By analyzing a set of elements as a group, FlexiCapture can increase accuracy by viewing elements "in context" and speeds up the steps needed to process the entire form.

A completed and tested FlexiLayout is easily exported into one of the ABBYY products - FormReader or FineReader Engine - for data capture of semi-structured documents, no matter how complex the layout may be.