Warning: Obsolete

RecoStarEngine was removed from DotImage in 11.0 and newer. This article is preserved for legacy reasons, but may go away in future updates. Please consider one of our other OCR engines for new development

RecoStar Engine

The RecoStar engine from OpenText is a hybrid engine that can recognize machine printed text as well as hand printed text.

Features

In addition to standard preprocessing options, the RecoStar engine includes tools for hole punch removal, line removal and shading removal.

Supported Languages

The RecoStar engine has different levels of support for languages. Languages that are Latin based will typically include a subset of the WinLatin1 character set. Some languages like Russian will include Cyrillic support in addition to Latin characters. Some languages like Japanese, Chinese and Thai support Arabic numerals, but not larger character sets such as Hanzi or Kanji. In most OCR engines, languages are associated with specific countries or locales. The RecoStar engine provides support for some broader categories such as "Western European". In these cases, a specific CultureInfo object is chosen to represent that larger region.

The default CultureInfo is English (en, 0x0009), which is equivalent to RecoStar's "Wester Europe" which includes a number of characters that have diacritical marks.

The CultureInfo German (de, 0x007) represents RecoStar's "Central Europe" (while the CultureInfo German - Germany (de-DE, 0x407) is specifically for Germany.

Supported Output Formatters

The standard DotImage output formatters Text and PDF will work with the RecoStar engine.

Deployment

To deploy the RecoStar engine, the client must negotiate a distribution license with OpenText. This license will typically be a file that can be passed to the RecoStar engine for verification. This license depends on the client's assembly being tied strong named and signed. The details of this process are handled by OpenText, not Atalasoft.

To run, the engine will need the folder OcrResources\RecoStar\x.y (where x.y corresponds to the current version shipping with DotImage) and all of its contents.

Example

The RecoStar Engine is used in exactly the same way as the other OCR engines, all of which inherit from the same base class, Atalasoft.dotImage.OCR.OcrEngine.

Special Considerations

The RecoStar engine is licensed in three ways: evaluation, development and deployment. Under evaluation, the engine is licensed for all features and runs no faster than 25 characters per second. Under development, each engineer must use a hardware dongle (typically a USB key), which will also limit speed and features. When deployed, the application will supply a distribution license file that will be tied to the client's assembly.

See Also

OCREngine

GlyphReaderEngine

TesseractEngine (retired in 11.1 - use Tesseract3Engine)

Original Article:

Q10365 - INFO: RecoStarEngine - Overview

INFO: RecoStarEngine - Overview

Atalasoft Knowledge Base
https://www.atalasoft.com/kb2/KB/50122/INFO-RecoStarEngine-Overview