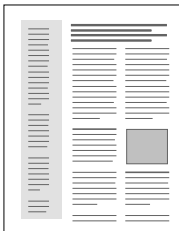
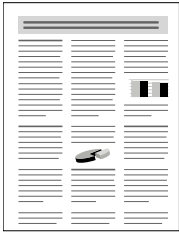


Exegenix technology interprets the structure of documents like humans do.



Exegenix Technology Overview

Exegenix Conversion Solutions convert into XML any file that can be printed to PostScript™ or PDF. Our extensible, modular technology is easily adapted to meet the content conversion needs of any organization, and minimizes human involvement via a combination of automated processing and intuitive interfaces that streamline conversion management.

Unique approach

Exegenix's revolutionary approach uses visual cues to uncover a document's structure, much the same way that humans do. People rarely have problems determining the hierarchical structure of any document they encounter, because they look at a document as a whole, taking into consideration each graphical object's format, position, and context. Exegenix technology does the same thing – it interprets a document's logical structure based on the appearance and position of its components.

Deployment flexibility

The modular structure of Exegenix technology means that the conversion process can be tailored to your unique workflow requirements. Deployment flexibility is achieved via the following components:

ECS Engine

The automated process that converts documents individually from PDF/PostScript to XML.

ECS Inspector

The intuitive interface that enables users at any technical level of expertise to take control of the conversion process via user-directed actions, if required. Authors can examine structures and text identified by the ECS Engine, and add value to the output XML prior to export.

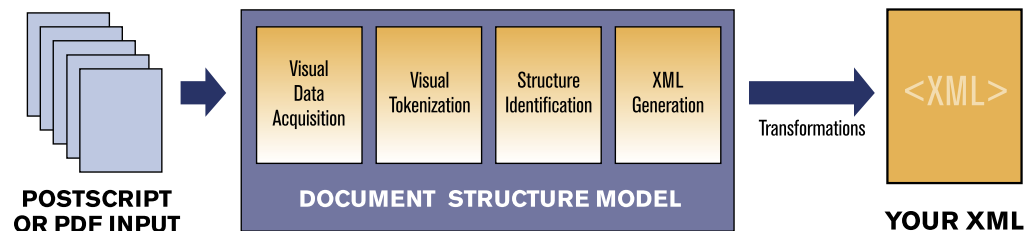
ECS Administration

A versatile configuration module adapted to each organization's specific requirements, forming the basis of a flexible workflow framework.

ECS Engine

All content goes through four processes within the ECS Engine in order to uncover the document's structure and generate valid XML that can then be transformed to meet specific customer needs.

Exegenix Technology's Internal Conversion Process



Visual Data Acquisition

The first step in the process analyzes a document's PostScript or PDF representation to extract all information about the appearance of the document. This includes the characters in the document and their typography, and any other visual objects. Because the process extracts text directly from the input datastream, all content is accurately retained during conversion.

Visual Tokenization

The Visual Tokenization phase identifies the basic building blocks of document structure, including many important visual cues, and the large-scale layout areas of the page.

Structure Identification

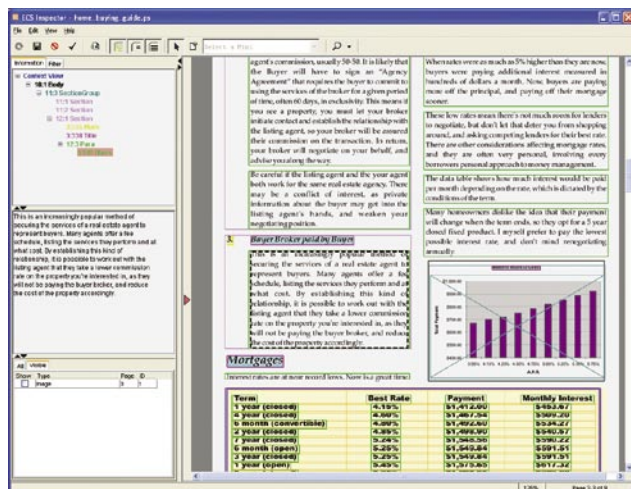
The Structure Identification phase places these basic building blocks into a tree structure. This phase identifies sections, paragraphs, quotes, lists, tables, footnotes, and other graphical objects, and forms a complete, cohesive, internal representation of the structured document.

XML Generation

This phase uses the internal representation of the document to export an XML file that not only presents the document's content and logical structure but also retains all relevant formatting information. The XML file is designed for ease of use in XSL Transformation scripts.

ECS Inspector

The ECS Inspector offers a unique preview into the conversion process, enabling further streamlining via user-directed action if required, and allowing the user to add value to the converted content, or



ECS Inspector: input document, with structures

specify how ambiguous input should be structured. Before any XML is generated, a user can examine the structures and text that the ECS Engine has automatically identified. Results are presented in an intuitive single window, so there is no need to open the original document in the application it was created with, and compare it with XML output. No XML expertise is required – the ECS Inspector is intuitive enough for even a novice user.

Added control

When documents are too complex for a human to easily understand, they will prove challenging for the ECS Engine also. Once the automated conversion process is completed, the ECS Inspector can be used to examine the resulting structures, and, via our “hinting” mechanisms, the ECS Engine can be directed to choose another path for recognition, e.g., turning a very complex table into a diagram, or considering a number of distinct embedded images as one single image.

Added efficiency

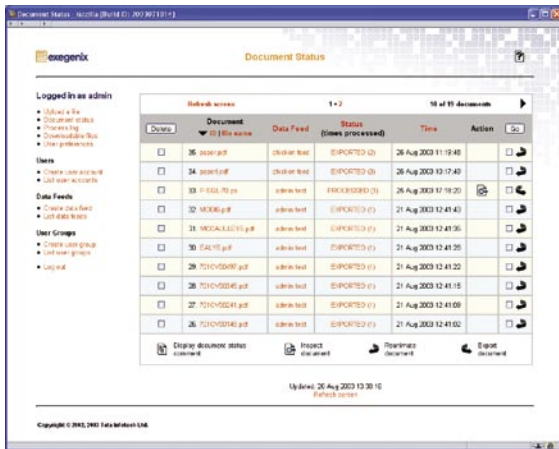
The ECS Inspector displays a representation of the original input document, with structures overlaid as outlines. This integrated interface is a practical and effective way for novices and XML experts alike to examine the structures discovered by the ECS Engine – a far more efficient process than the effort required to examine markup structures in an XML editor, while simultaneously referring to the original input document.

Added value

The ECS Inspector also provides extensive value-add capabilities, via point-and-click classification of objects' roles. In a financial report, for example, text objects could quickly and easily be differentiated as “Corporate Profile” or “Management Bio”, tables as “YTD Revenue” or “Projected Production Output”, images as “Corporate Logo” or “Product Image”, and charts as “Market Share” or “Month to Month Revenues”. The possibilities are endless, as are the post processing capabilities afforded by this elegant, point-and-click interface.

ECS Administration

The ECS Administration interface enables each organization to manage its specific conversion workflow processes. Via a standard Web interface, and based on user privileges, conversion teams can upload documents (individually or in batches), view the status of documents at various stages of processing, filter their view to optimize their productivity, and approve documents for final XML export.



ECS Administration: on demand, via browser

Ready to use, out of the gate

Other conversion technologies provide little support to help you get the software up and running: you must install the software on your own servers; configure the software to work with your other processes; develop systems and processes to manage input and output material, and monitor productivity and workflow; and provide access mechanisms to your users. In comparison, ECS Administration interfaces support multiple users, using a built-in “workgroup” or “project” based paradigm. You can organize files into distinct datafeeds, filter and sort conversion jobs, and examine the conversion process at every step, as required. With no systems to build, performance monitoring and cost analysis reporting are straightforward.

Lightweight requirements, heavyweight utility

ECS Administration interfaces are all browser-based, and deliver an easy-to-understand, familiar interface to all your conversion operators via standard HTML and CSS – anywhere in the world. The system is on secure server software, using secure data transmission protocols that are proven to keep your data well-protected. Our wide variety of deployment options makes it easy to use and administer our conversion technology, no matter what the size or nature of your organization.

Exegenix Technology: Features and benefits summary

ECS Engine

- Feature:** The process employs human-like intelligence to interpret each page.
- Benefit:** Structures are accurately converted automatically, with no scripting required.
- Feature:** The process is based on a document’s presentation, not binary data files.
- Benefit:** Conversion is not subject to inconsistent

application of original styling/formatting codes, and resulting structural errors.

- Feature:** The process extracts text directly from the input datastream.
- Benefit:** All content is accurately retained during conversion.
- Feature:** Conversion is based on easily generated, well documented, and accurate page description formats: PostScript and PDF.
- Benefit:** Consolidate your conversion processes around a single technology, with no need to maintain software from multiple vendors for individual file formats. Configuration is not required for every different input document type.

ECS Inspector

- Feature:** Presents an integrated view of both input and output.
- Benefit:** Any conversion review process is streamlined and efficient.
- Feature:** No XML editing or direct manual input.
- Benefit:** Both the conversion skillset and the XML tool requirements are lowered.
- Feature:** Object boundaries and types can be indicated via a flexible point-and-click “hinting” process.
- Benefit:** Any errors resulting from ambiguous input can be quickly corrected.
- Feature:** Hints can be generalized for use across multiple input documents.
- Benefit:** Adding value to all documents in an input datastream enhances both manual and automated downstream processing.
- Feature:** Individual documents can be annotated.
- Benefit:** Downstream processing of a specific document is enhanced.

ECS Administration

- Feature:** Browser-based.
- Benefit:** Any computer with a web browser can access ECS. Computer system requirements are lightweight; if the computer is capable of surfing the web, it can interact with ECS. Interface is familiar to users.
- Feature:** Standards-compliant back end.
- Benefit:** Based on tested, proven, secure server technologies. Supports multiple users with ease. Scales effortlessly.
- Feature:** Built-in users/user groups/datafeed model.
- Benefit:** Interface delivered on demand via browser, with user-specific options and data access.

Exegenix shatters old models

Until now, organizations needing to convert their existing content into XML have had three options:

Manual conversion

When content is tagged by hand it's easy to introduce errors into the conversion process. Significant post-conversion review and editing will generally be required before you can use your XML. Manual conversion is also extremely resource-intensive and can involve a lengthy turnaround time.

Compared to manual conversion:

Exegenix conversion technology is accurate, fast, and minimizes human resource requirements enabling valuable expertise to be employed more effectively elsewhere.

“Scripted” conversion

Scripting software relies on rigid conformance to formatting styles in the original word processor document. However, human authors generally do not create documents in such a consistent manner, and low-quality XML output can result. These “solutions” can also involve costly and lengthy custom programming for each type of document, and for each proprietary application that is used to create those documents.

Compared to “scripted” conversion:

Exegenix technology removes the dependency on consistently-applied formatting styles, because it interprets a document's logical structure based on the appearance and position of its components in the same way a human reader does. It does not require programming expertise to develop and maintain configuration scripts every time a format changes, every time the source data changes, or every time the authoring product version changes.

“Save as XML”

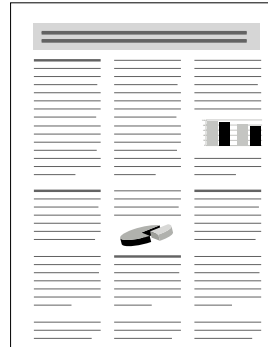
XML exported from a word-processing program has flat structure, and is formatting-based. The many combinations of formatting required to signify specific output structures must be known in advance and rigidly adhered to by the document author, and configuration or transformation scripts must be developed for every distinctly formatted input dataset.

Compared to “Save as XML”:

Exegenix technology analyzes each document's typographical characteristics (such as predominant font, largest type, indented, superscripted), discovering structures and the purpose of objects (Document Title, Subtitles, Paragraphs, Lists) using an analysis of the entire document.

Revolutionary technology, superior XML

Exegenix technology takes you from something like this:



To something like this:

```
<document><body>
  <section id="E23-2" level="1">
    <beginpage pagenumber="1"></beginpage>
    <title id="E4-71"><line id="I4">Buying Your
  </line><line id="I5">First Home</line></title>
    <sectionbody>
      <bridgehead role="subtitle">
        <line id="I6">A guide for first-time home</line>
        <line id="I7">buyers purchasing a</line>
        <line id="I8">resale property</line>
      </bridgehead>
      <para id="E11-9"><paraseg><block id="E1-270">
        <specialcap id="E1-362" font-size="49.3pt"
        font-weight="normal">H</specialcap>aving recently
        gone through the exhilarating and tiring
        experience of buying my first home, a resale
        high-rise condo property, I certainly learned a
        great deal about a number of home-buying issues.
      </block></paraseg>
      </para>
    </sectionbody>
  </section>
</body></document>
```