

# Exsys Data Extractor (EDX)

## Features & Benefits

The challenge is “time to market” and the need for automatic and accurate extraction of key market data from multiple Internet sources and publishing that data timeously on various target systems. Lack of content can mean potential lost revenue and disappointed paying customers. It’s a competitive world and being responsive to customer needs is the name of the game.

What if there was a way that process could be improved by many factors and backlogs for collating data could be drastically shortened ? Customers would see a marked improvement in content collection, availability, efficiency, reliability and quality.

Exsys Data Extractor (EDX) can give you that capability – now.

Here is a list of many key features and benefits of the software.



### Point-and-Click Rule Creation

Easily create extraction rules simply by browsing websites – no coding required.



### Automatic Download of Source Data

Download, according to user defined schedule and frequency, HTML, PDF, XML (or other) files from Internet or Intranet.



### Specify Numerical or Textual data

Extract only the precise content you need utilizing a wide variety of user defined rules.



#### Poll live HTML pages

Any frequency of polling to live pages to check for changes. Dramatically reduce downstream processing by only delivering content that has changed.



#### Easily build Thomson Reuters contribution messages

Contributions can be to any permissioned RIC/FID combination using either Exsys DataBridge software or data can be sent to other target systems. EDX can contribute messages to any target system or database.



#### Prevent unnecessary contributions

EDX will detect if data has changed or not and avoid unnecessary contributions (forced contribution to change date time available)



#### Fine-Grained Scheduling

Schedule extractions to match source update rates of critical content. These can be set for specific day/s of the week and specific time windows. EDX can automatically extract from web pages and create schedules based on published Holidays. Active day files can be used to force activity on special pre-defined dates.



#### Multi-source Extraction

Easily extract content that spans multiple pages and contribute to multiple RICs.



### Intelligent website management

EDX adapts to most website changes using intelligent word recognition



### Language-Agnostic

Extract content from sites in any language



### PDF, HTML, XML Content Retrieval

Automatically download specific data from most formats and other files from websites.



### Intelligent Site Navigation

Better extraction performance, as well as lower footprints on sites being extracted



### Content Filtering

Carry out precise searches for content that matches specific keywords or numerics



### Point-and-Click Content Transformation

Easily preprocess content into the exact format needed for Thomson Reuters field rules.



### Content Normalization

Reduce downstream processing by normalizing content to a common schema.



### Change Highlighting and Alerts

Instantly alert users when tasks have either succeeded or failed by sending emails or SMS with appropriate highlights.



### Multiple Delivery Formats

Deliver data in the format that you need it, including contribution messages, databases (Deliver data directly to SQL Server, any ODBC-compliant), other applications via TCP, HTML or other custom targets using Exsys DataBridge software.



### Logs

EDX maintains a complete audit trail of events.

Talk to Exsys about a demo of EDX and our roadmap for future enhancements and let us know if you have any specific requirements. [Click here for more information.](#)