

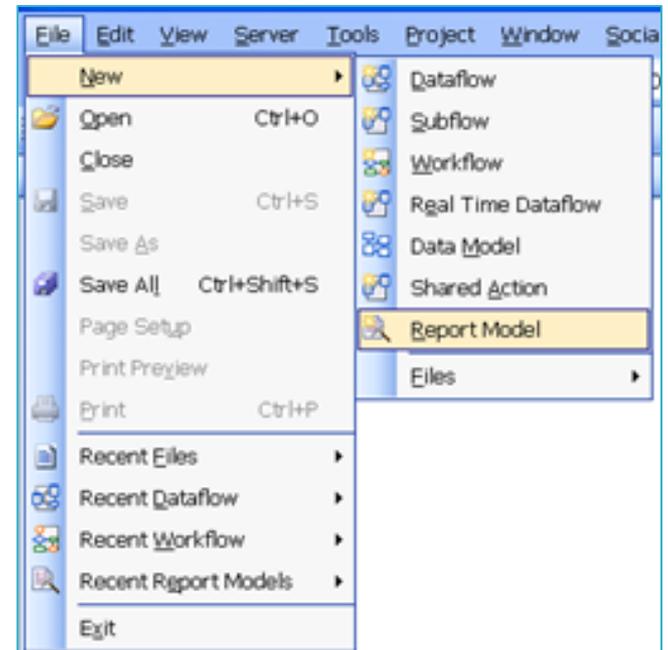
ReportMiner

Extracts Business Data From Printed Documents



Overview

- Information drives today's businesses and manually transferring unstructured data into electronic formats is no longer a cost-effective option
- ReportMiner liberates data trapped in printed documents such as reports, bank statements, and invoices in popular formats such as PDF, PRN, CSV, XML, and text-based files and
- Converts the data for use in operational and business intelligence applications

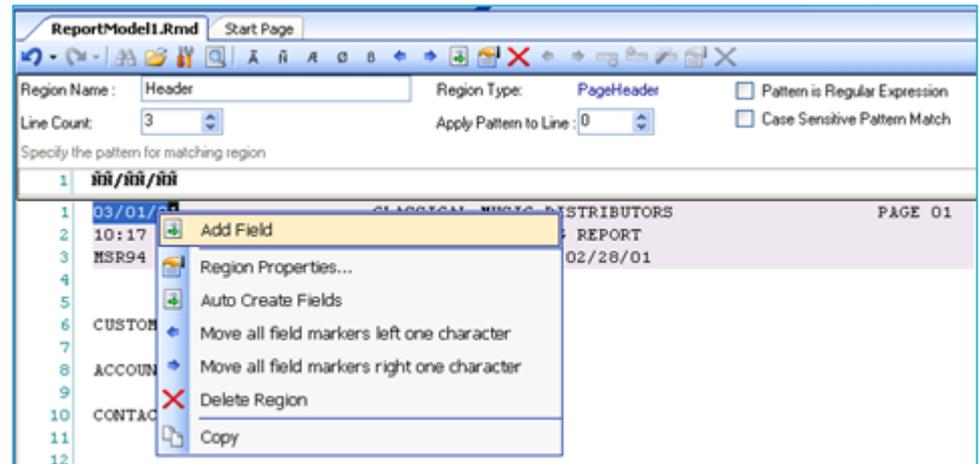


Key Features

- Intuitive user interface
- Hierarchical data extraction
- Single-click preview
- Data mapping and export to any supported destination and format
- Report models can be saved and used over and over again for subsequent conversions, saving even more resource time and money
- Performance: multi-core and multiprocessor hardware mine and deliver data quickly and efficiently

Intuitive User Interface

- WYSISYG user interface enables users to view the document and define regions and fields
- Auto Create feature highlights all occurrences of the regions and fields and instantly identifies all the fields in a region and determines their data types and sizes
- This capability enables users to define and verify a report model in minutes



Hierarchical Data Extraction

- Data extraction can be done from complex hierarchical reports in a single pass using the hierarchical model capability of ReportMiner
- The Auto Create feature scans the document and automatically adds the fields

The screenshot displays the 'Report Browser' window with a hierarchical tree view on the left and a 'ReportModelLRend' window on the right. The tree view shows a 'Record' structure with 'Header', 'CustomerData', and 'OrderData' sections, each containing various fields like 'PageNumber', 'OrderDate', 'CompanyName', etc. The 'ReportModelLRend' window shows a preview of the extracted data, including a table with columns for 'MEDIA QTY', 'DESCRIPTION', 'LABEL/NO.', 'UNT_PRC', and 'AMOUNT'.

| MEDIA QTY | DESCRIPTION | LABEL/NO. | UNT_PRC | AMOUNT |
|-----------|--------------------------------------|------------|---------|--------|
| 6 | Stravinsky, Dumbarton Oaks Concerto | SMCD 5120 | 8.99 | 53.94 |
| 1 | Schubert, Sonata in e, D.566 | AS-325 | 9.00 | 9.00 |
| 3 | Mozart, Symphony No.23 in D | CO-77884 | 8.99 | 26.97 |
| 6 | Schoenberg, Ode to Napoleon | CHAN 9116 | 9.59 | 57.54 |
| 2 | Shostakovich, 24 Preludes for piano. | CDA 66620 | 5.39 | 10.78 |
| 9 | Balakirev, Symphony no. 1 | ENTPD 4110 | 9.59 | 86.31 |
| 5 | Holst, St. Paul's Suite for Orch. | CBT-1020 | 5.99 | 29.95 |

Hierarchical Data Extraction

- Alternatively, users can manually add a field by highlighting the field in the object with the mouse cursor, right-clicking it, and selecting Add Field
- A new field is added to the Report Browser and the Report Definition Editor shows all the occurrences of this field in the report

The screenshot displays two windows from the Astera software. The left window, 'Report Browser', shows a tree view of report objects. The right window, 'ReportModelLRend', shows the report definition editor with a preview of the report output.

Report Browser Tree View:

- Record
 - Header
 - PageNumber (Integer)
 - OrderDate (Date)
 - CompanyName (String)
 - FromDate (Date)
 - ToDate (Date)
 - CustomerData
 - CustomerName (String)
 - AccountNo (Integer)
 - ContactName (String)
 - OrderData
 - Field_0 (Integer)
 - Field_1 (Date)
 - OrderDetails
 - Field_0 (String)
 - Field_1 (Integer)
 - Field_2 (String)
 - Field_3 (String)
 - Field_4 (Real)
 - Field_5 (String)

Report Definition Editor Preview:

Region Name: OrderDetails Region Type: Data
 Line Count: 1
 Specify the pattern for matching region: `^`

| 1 | CLASSICAL MUSIC DISTRIBUTORS | | | | PAGE 01 |
|--|------------------------------|--------------------------------------|------------|---------|---------|
| 10:17 | MONTHLY SHIPPING REPORT | | | | |
| MSR94 | FROM 02/01/01 TO 02/28/01 | | | | |
| 6 | CUSTOMER: Big Shanty Music | | | | |
| 8 | ACCOUNT NUMBER: 17959 | | | | |
| 10 | CONTACT: Marvin Mabry | | | | |
| 13 | MEDIA QTY | DESCRIPTION | LABEL/NO. | UNT_PRC | AMOUNT |
| ORDER NUMBER: 536118 SHIP DATE: 02/04/01 | | | | | |
| CD | 6 | Stravinsky, Dumbarton Oaks Concerto | SMCD 5120 | 8.99 | 53.94 |
| | 1 | Schubert, Sonata in e, D.566 | AS-325 | 9.00 | 9.00 |
| | 3 | Mozart, Symphony No.23 in D | CO-77884 | 8.99 | 26.97 |
| | 6 | Schoenberg, Ode to Napoleon | CHAN 9116 | 9.59 | 57.54 |
| ORDER NUMBER: 536131 SHIP DATE: 02/13/01 | | | | | |
| CD | 2 | Shostakovich, 24 Preludes for piano. | CDA 66620 | 5.39 | 10.78 |
| MD | 9 | Balakirev, Symphony no. 1 | ENTPD 4110 | 9.59 | 86.31 |
| TAPE | 5 | Holst, St. Paul's Suite for Orch. | CBY-1020 | 5.99 | 29.95 |
| ----- | | | | | |
| 03/01/01 | CLASSICAL MUSIC DISTRIBUTORS | | | | PAGE 02 |
| 10:17 | MONTHLY SHIPPING REPORT | | | | |
| MSR94 | FROM 02/01/01 TO 02/28/01 | | | | |
| CUSTOMER: Classics and Jazz | | | | | |

Single-Click Preview

- Single-click preview capability verifies and tests the report model as it is being built. This preview shows extracted data and any conversion or validation errors.

Data Preview

Source Record Count: 50

Data Preview for action ReportModel1. Total Records: 13. Records With Errors: 0. Duration: 00:00:00

| Object Path | PageNumber | OrderDate | CompanyName | FromDate | ToDate |
|-------------|------------|-------------------|---------------|-------------------|--------------------|
| Header | 2 | 3/1/2001 12:00:00 | CLASSICAL MUS | 2/1/2001 12:00:00 | 2/28/2001 12:00:00 |

| Object Path | CustomerName | AccountNo | ContactName |
|--------------|------------------|-----------|-------------|
| CustomerData | Big Sharly Music | 17959 | Marvin Maby |

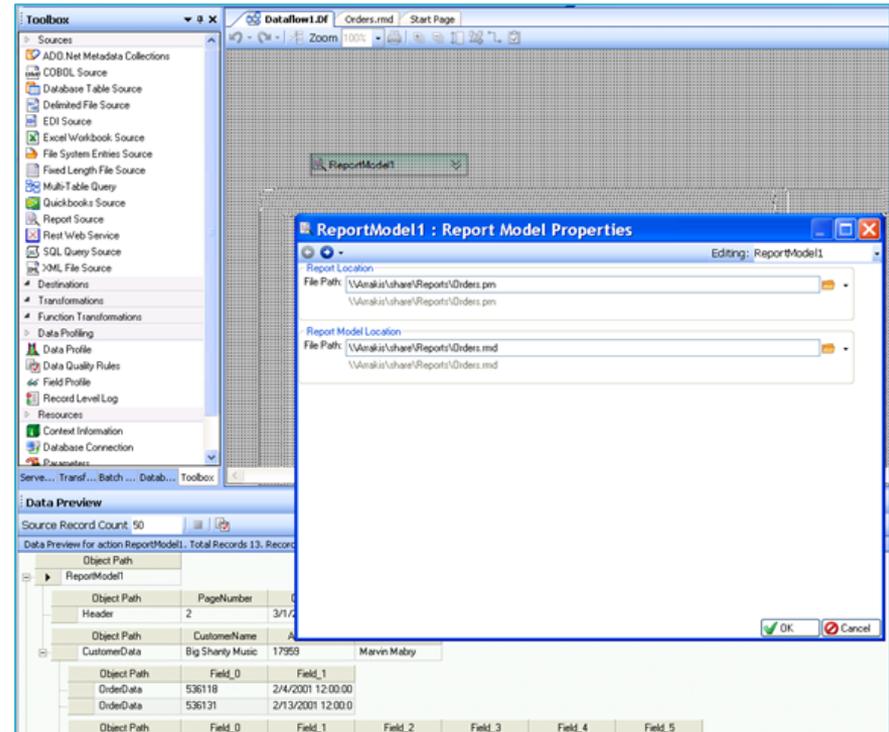
| Object Path | Field_0 | Field_1 |
|-------------|---------|--------------------|
| OrderData | 536118 | 2/4/2001 12:00:00 |
| OrderData | 536131 | 2/13/2001 12:00:00 |

| Object Path | Field_0 | Field_1 | Field_2 | Field_3 | Field_4 | Field_5 |
|--------------|---------|---------|---------------------|------------|---------|---------|
| OrderDetails | CD | 6 | Stravinsky, Dumb | SMCD 5120 | 8.99 | 53.94 |
| OrderDetails | | 1 | Schubert, Sonata | AS-325 | 9 | 9.00 |
| OrderDetails | | 3 | Mozart, Symphon | CD-77884 | 8.99 | 26.97 |
| OrderDetails | | 6 | Schoenberg, Ode | CHAN 9116 | 9.59 | 57.54 |
| OrderDetails | CD | 2 | Shostakovich, 24 | CDA 66620 | 5.39 | 10.78 |
| OrderDetails | MD | 9 | Balakirev, Symph | ENTPD 4110 | 9.59 | 86.31 |
| OrderDetails | TAPE | 5 | Holst, St. Paul's S | CBT-1020 | 5.99 | 29.95 |

Job Progress | Verify | Data Preview | Quick Profile

Data Mapping and Export

- Data mapping and export to any supported destination and format
- Supported destinations include databases like SQL Server, Access, MySQL, and any ODBC compatible database; file formats such as fixed length, delimited, Excel, and XML



Re-Use for Subsequent Conversions

- Report models can be saved and used over and over again for subsequent conversions
- This saves even more resource time and money

Performance

- Designed to efficiently process large data volumes
- Built from the ground up as a high-performance parallel-processing engine
- Takes full advantage of today's multi-core and multiprocessor hardware to mine and deliver data quickly and efficiently

ReportMiner Advantages

- Flexible and intuitive visual interface: enables users with little or technical background to accomplish data extraction without employing expensive IT resources
- Hierarchical data extraction: enables extraction of complex data
- Data mapping and export: any supported destination and format
- Re-Use: report models can be saved and reused
- Performance: multi-core and multiprocessor hardware mine and deliver data quickly and efficiently