REQUEST FOR PROPOSAL
DATA WAREHOUSING
Extraction Transformation and Loading (ETL) TOOL
FOR
RENSSELAER POLYTECHNIC INSTITUTE

Rensselaer Polytechnic Institute
Troy, New York
http://www.rpi.edu
October 2001
1 **Objective**

This Request For Proposal (RFP) is issued for the purpose of supporting the institute-wide initiative of developing data warehousing architecture to provide better access to information to foster better and more informed decision-making. The ETL tool is an integral part of the data warehousing architecture that will be used by the data warehousing professional staff to model and create target warehouse, extract data from multiple data sources, transform the data to make it accessible to business analysis, and loading multiple target data marts.

2 **About Rensselaer**

Rensselaer has been at the forefront of scientific and technological education since 1824. This year U.S News & World Report ranked Rensselaer as one of the top 50 national universities. Its long-standing reputation for educational excellence now draws students from every state and more than 83 foreign countries.

Rensselaer is a private university located on a 260-acre campus in Troy, New York, and is comprised of five schools serving both undergraduate and graduate students. The faculty is comprised of approximately 360 members, and staff represents approximately 1,350 people. The student population is approximately 5,000 undergraduate and 1,800 graduate students.

3 **Response Requirements:** Please submit your response in eight (8) copies, tabbed as shown below.

4 **Executive Summary:** Include in this tab an executive summary discussing the highlights of the proposal. Do not include pricing information in this tab.

5 **Introduction:** Include in this tab any information your company wishes to submit about the nature of your business and primary business focus. Include information about what makes your company (not the product) different from your competitors.

6 **Proposal Format:**

6.1 Rensselaer will be imaging all proposals. DO NOT BIND the proposal. The official name of the firm submitting the proposal must appear on the outside front cover of each binder, together with the title “RFP Data Warehousing ETL Tool Proposal”.

6.2 Each Proposal page must be numbered consecutively from the beginning of the proposal (Executive Summary) through all appended material.

7 **Contract Terms:** Include copies of all contracts or agreements you expect Rensselaer to sign.
8 **Exceptions:** Include in this section either a statement that you will make no exceptions to the requirements of the RFP, or a statement clearly indicating any exceptions and a statement of substitute wording for resolving the exception.

9 **Company Profile:** Please include the following information about your company; Nature of business and primary business focus, percentage of business that comes from tool referenced in the RFP, major company differentiator – what makes your company (not product) different, address for headquarters, address for office closest to Rensselaer (Troy, NY), contact information including - name, title, address, phone, fax, email and web site address.

10 **Financial Stability** - Each proposal must include a certified external audit statement and the 10-K report of the last corporate fiscal period for the firm submitting the proposal. Each Vendor must submit documentation indicating at least three years of experience, as of the Proposal submission date, in providing services similar to those required in this RFP.

11 **Billing Methods:** Describe in detail your billing methodology. Rensselaer is exempt from payment of certain taxes. It is expected that invoices will not include any taxes for which an exemption applies. A tax exemption certificate can be provided upon request.

12 **Costs:** Please specify pricing for each component. Indicate if pricing would differ if only certain parts of the system were purchased by Rensselaer.

13 **Support:** Proposals must indicate the name, experience level, and length of service with the firm of the customer service representative who would be assigned to Rensselaer. A description of the on-going support program must be provided. Describe your data mart development methodology. Describe levels and hours of support available.

14 **References:** Please list at least five (5) clients that you have done business with in the past year. Include the company's name, address, telephone number, contact name and number of years as a customer.

15 **General Product Requirements:** For the next three years Rensselaer embarks on the institute-wide initiative of developing a data warehousing architecture to provide better access to information to foster better and more informed decision-making. The ETL tool is an integral part of the data warehousing architecture that will be used by the data warehousing professional staff to model and create a target warehouse, extract data from multiple data sources, transform the data to make it accessible to business analysts, and load multiple target data marts. Our primary ERP system is Banner from SCT. The SCT Banner product at Rensselaer currently resides on IBM RS6000 machines (32 bit upgraded to 64 bit architecture), running AIX 4.3.x with Oracle 7.3.x (to be upgraded to Oracle 8.1.7.x).
Detailed System Requirements

15.1 System Requirements are described in Appendix A.
15.2 Selection Criteria A team from the IACS and Purchasing Departments will evaluate the proposals. Clear and concise responses are requested. After an initial review, this team may determine that further discussions and/or vendor presentations may be required. The following criteria (not listed in order of priority) will be used to select the firm:

15.2.1 The degree to which the proposed system satisfied the Rensselaer business requirements and management process.
15.2.2 Overall system functionality and flexibility, including ease of use
15.2.3 Initial cost of product and ongoing license and maintenance costs
15.2.4 Support and training provided for implementation
15.2.5 Corporate financial condition
15.2.6 Ongoing maintenance and training
15.2.7 Ability to accept Rensselaer’s terms and conditions
15.2.8 References

16 Notification of Award/Contract: Rensselaer will notify the successful vendor through a letter of intent. A contract will be negotiated between Rensselaer and the successful vendor. The contract will, among other provisions, incorporate this RFP and the successful bidder’s proposal. Upon execution by the Vendor and Rensselaer, the contract will be submitted for final approval and a Purchase Order will be issued. Rensselaer will notify unsuccessful vendors in writing.

17 Rensselaer reserves the right to:

17.1 Reject any and all proposals received in response to this RFP.
17.2 Waive or modify minor irregularities in proposals received, after prior notification to the Vendor.
17.3 Adjust or correct cost or cost figures with the concurrence of the Vendor if errors exist, and the Vendor establishes that a verifiable error occurred in the computation of the proposal.
17.4 Adopt all or any part of a bidder's proposal in selecting the optimum configuration.
17.5 Negotiate with selected Vendor responding to this RFP within the RFP requirements necessary to serve the best interests of the College.
17.6 Begin contract negotiations with another Vendor in order to serve and realize the best interests of the College, should the College be unsuccessful in negotiating a contract with the selected Vendor within an acceptable time frame.

18 Liability: The Institute is not liable for any costs incurred by a Vendor in the preparation and production of a proposal or for any work performed prior to the issuance of a contract or delivery order.

19 Disaster Recovery Responsibility: The Vendor must assume an active facilitating role in any major system failure. They must act as the primary source of technical expertise relative to the rapid re-establishment the system.
20  **Project Schedule:**

<table>
<thead>
<tr>
<th>Event</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Issue RFP</td>
<td>October 03, 2001</td>
</tr>
<tr>
<td>RFP response due date</td>
<td>October 17, 2001</td>
</tr>
<tr>
<td>Vendor presentation</td>
<td>October 22-26, 2001</td>
</tr>
<tr>
<td>Decision date</td>
<td>November 02, 2001</td>
</tr>
</tbody>
</table>

**General Information:** Interested firms must respond to this Request for Proposal in order to be considered. Questions regarding this Request For Proposal must be addressed to Ora Fish, Project Manager, IACS Phone (518-276-2213).

Interested vendors must submit an original and (8) eight copies of the proposal, in a sealed envelope clearly marked Data Warehousing ETL proposal to:

Keith E. Martens  
Purchasing Agent  
Purchasing Department – West Hall 311  
Rensselaer Polytechnic Institute  
110 8th Street  
Troy, NY 12180-3590

Electronic submission must be sent via email to: [martek@rpi.edu](mailto:martek@rpi.edu)  
If hand carried, the office hours are 8:30 a.m. – 5:00 p.m. Facsimile responses will **NOT** be accepted.

Proposal material will be treated as proprietary and become the property of Rensselaer. Rensselaer reserves the right to waive any irregularities in the proposals and to accept or reject any or all proposals.
General Terms and Conditions

20.1 If this Request for Proposal results in an agreement to provide a data warehousing ETL tool, the following terms and conditions will apply:

20.2 This Agreement can be modified only in writing and when signed by both parties.

20.3 Either party may terminate this Agreement, in part or in whole, without penalty.

20.4 Upon notification of termination, the two parties shall agree upon a turnover plan and agree upon the compensation due.

20.5 This entire Agreement may not be assigned, sublet, or transferred without the prior written consent of Rensselaer.

20.6 Vendor is an independent contractor, not an employee, agent or partner of Rensselaer. Therefore, neither Vendor nor any of its employees are entitled to participate in any form of benefit or privilege that Rensselaer extends or may offer to any of its own employees.

20.7 Vendor agrees to indemnify Rensselaer and to hold Rensselaer harmless from and against all claims, liability, loss, damage, and expenses (including legal fees), arising from or due to any claim with respect to any part of the sales or services covered by this Agreement or any activity, Vendor, its officers, agents, or employees on or about Rensselaer’s property. Vendor shall defend any such litigation brought against Rensselaer. This clause shall survive termination of this Agreement.

20.8 During the performance of this Agreement, Vendor agrees not to discriminate against any individual because of race, color, religion, sex, or national origin, or because he or she has a physical or mental handicap or because he or she is a disabled veteran or a veteran of the Vietnam era. The aforesaid provision shall include, but not be limited to, the following: employment; upgrading; demotion; transfer; recruitment or recruitment advertising, layoff or termination; rates of pay or other forms of compensation; and selection for training, including apprenticeship, sales, and conditions of consultations regarding special needs of customers and/or clients. Furthermore, vendor will adhere to rigorously enforced principles of affirmative action regarding members of minority groups and handicapped individuals.

20.9 Services performed under this Agreement must be in accordance with all governmental laws, rules, regulations, and ordinances, including, but not limited to, OSHA, ANSI, EOA, ENCON, and Vendor certifies to this requirement. Rensselaer reserves the right, at its sole option, to order cessation of performance in case of violation by Vendor, and Rensselaer shall have no liability whatsoever resulting from such interruption.

20.10 Parking permits may be required for parking on campus. Rules issued by the Parking Office must be followed. If any fines are imposed on Vendor personnel, it is the responsibility of Vendor to appeal the violation and/or pay the fine.

20.11 This Agreement shall be governed by and construed in accordance with the laws of the State of New York.
21 Insurance Requirements

Before any services and/or work can be performed on Rensselaer’s premises, evidence of insurance in force naming Rensselaer as an additional insured must be in the possession of the Rensselaer Department of Risk Management, 110 8th Street, Troy, New York 12180. Unless otherwise directed in writing, the following coverage’s are required:

Comprehensive General Liability (including operations and completed operations) - $2,000,000 – occurrence, $2,000,000 – aggregate.
Comprehensive Automobile Liability (including owned, non-owned and hired autos) - $1,000,000 combined single limit.
Workers Compensation as required by law.

The vendor is responsible to maintain insurance coverage throughout the term of the agreement and/or contract. If for any reason during the term, the insurance policy is cancelled the consultant must immediately notify Rensselaer.
Appendix A

**Product Profile:** Provide a one-page product profile for each product used in the response to this RFP. Please include the following:

1. Product name
2. Product description
3. Current release level
4. Date current release level was generally available
5. Projected general availability of next release level
6. Current product install base
7. Number of companies
8. Number of users

**Consulting:** Do you employ a refined, tried-and-tested data mart development methodology? How experienced are the individual consultants?

**Extraction, Transformation, and Loading functionality:** Provide a complete description of your proposed ETL solution. Including the key features of your ETL solution and products, as well as what makes your solution unique.

Please address the following specific questions:

**Completeness**

1. Provide a complete description of all phases of the data warehouse population that your proposed solution addresses: Modeling, extraction, transformation, loading, managing the data repositories and warehouse administration.
2. Describe the design/development tool set. Point and click environment versus programming.
Development Environment

Please address the following as it pertains to your product:

1. Ease of Use. How quickly can developers get up to speed with the tools, and how much support will they require?
2. Requires knowledge of 4GL
3. Support complete development environment, including versioning and run-time debugger
4. Ease of promoting transformation from development to production
5. Version/configuration management.
6. Easily debug transformation logic
7. Re-usable functions and automatic propagation of changes
8. Dependency analysis (assess impacts of changes). Dependency and analysis feature of source data changes
9. Specification of ETL functions using pre-packaged transformation objects, accessible via an intuitive graphical user interface
10. Ability to specify complex transformations using only built-in transformation objects. The goal is to specify transformations without writing any procedural code
11. Incremental aggregation and computation of aggregates by the ETL tool in one pass of the source data
12. Automated, slowly changing dimension support (Type I, Type II, Type III)
13. Able to join data from multiple sources. Support for concurrent processing of multiple source data streams, without writing procedural code.
14. Reuse individual transformations
15. Reroute bad records to separate target
16. Implement conditional logic for updates
17. No requirement to generate and compile source code
18. No requirement for intermediate disc files
19. Can the tool exploit outside cleansing and transformation routines and allow developers to embed their own transformation functions into it
20. Team Development. Does the toolset allow multiple developers to work on the same project concurrently, sharing ideas and results? Can the developers attach to the development environment via a LAN, WAN, or the Internet?
21. Support for data extraction, cleansing, aggregation, reorganization, transformation, calculation, and load operations, including the following functions:
   o Filter data, convert codes, perform table lookups, calculate derived values
   o Validate data to check content and range of field values
   o Perform procedural data cleansing functions
   o Load cleansed data to the target data mart or central DW
   o Re-usage of the stored procedures and functions from our production Oracle based system
   o General transformation support (how many different transformation functions are available)
   o Automatic generation of sequence numbers
Management
Please address the following as it pertains to your product:

1. Does the toolset schedule, coordinate, and execute all the steps involved in populating a data mart on a regular basis?
2. Describe data warehouse administration functions
3. Ease of installation
4. Ability to monitor and manage the runtime environment in real time
5. Support graphical job sequencer, and nesting of sessions
6. Produce audit and operational reports for each data load
7. Automatic generation of centralized metadata
8. Automatic generation of data extract programs
9. Support for the analysis of transformations that failed to be accepted by the ETL process
10. Extensive reporting of the results of an ETL session, including automatic notification of significant failures of the ETL process
11. Ability to schedule ETL sessions based on time or the occurrence of a specified event, including support for command-line scheduling using external scheduling programs
12. Ability to schedule FTP sessions based on time or event. (FTP remote flat file to Application server: UNIX or NT)
13. Access data from multiple, operational data sources: Native Interface to Oracle, Flat File interface, XML sources, Others sources
14. Restart on recovery (will the transformation process restart automatically from abnormal termination in a recovery mode, restart logic)

Scalability
Please address the following as it pertains to your product:

1. How scalable is the ETL tool? Can it allow scalability from 100MB data source to a 100GB data source without significantly degrading performance or requiring awkward database or platform swaps?
2. Distributed data warehouses / data marts
3. Parallel server engines (multiple application servers)
4. From one data source to multiple data sources

Platform extensibility.
1. Does the tool run on UNIX and NT
2. Describe options available to us in terms of moving the product from the NT environment to the Unix environment, from a product and price perspective.

Design and Modeling: Describe the functionality available to define the logical and the physical data models, and create indexes.
Performance

Please address the following as it pertains to your product:

1. Directly executable code that runs on multithreaded UNIX or NT server engine
2. One path for populating detailed transactions and aggregated data
3. Does the product require intermediate files to extract and transform relational sources or can all calculations be performed in high speed memory using multithreaded processes
4. How fast can the toolset extract, transform, and load data?
5. Supports tuning
6. In-memory data handling
7. Parallel execution of jobs

Meta Data

Please address the following as it pertains to your product:

1. Meta Data Exchange architecture. Generate, manage, and maintain a central metadata repository that contains:
   - Source data definitions
   - Target data models
   - Transformation rules
   - Derived computations
2. Does the metadata repository consist of both business and technical definitions, and can it be easily browsed by end users and power users via client/server and/or Web connections?
3. Automatic generation of central metadata, including source data definitions, transformation objects, target data models, and operational statistics
4. Metadata exchange architecture that supports automatic synchronization of central metadata with local metadata for multiple end-user BI tools
5. End-user access to central metadata repository via a right-click
6. Metadata exchange API compliant with COM, UML, and XML
7. Support of metadata standards, including OLE DB for OLAP

Security

1. Describe the available levels of security.
2. Encrypts or hides passwords or connect strings to databases
3. Allows source and target database security to prevail
4. Requires registration of users to gain access to environment
5. Defines enforceable roles for development
6. Defines enforceable roles for operations (session execution)
7. Defines enforceable roles for administration
8. Allows for user permissions to be set by work area (or folder)
Your Partners
Please indicate products that are fully integrated with your solution. Please describe the connectivity and the interface requirements.

1. BI tools
2. Budgeting and Financial Applications
3. CRM applications
4. Analytic applications
5. Corporate portals
6. Others

**Price and licensing model:** How much does the package cost, including up-front training, support and consulting?