Issue #4 • August 2004

Tree • times

A Publication of Treehouse Software, Inc.

This Issue

*"I want to extend my thanks for the support I received this last weekend from Treehouse Software employee **Lynn McIntyre**. She spent hours on the phone with me as we attempted to resolve an error that was occurring in **N2O**.*

She pointed me to a possible problem with the NATURAL parameters. It turned out that somehow the parms got changed from 1200 seconds to 2 seconds on the TT NATPARM, and if it hadn't been discovered and resolved before Monday Morning, it would have shut down nearly our entire system and hundreds of people would not have been able to do their work. Lynn really went above and beyond.

I've also received excellent support from the entire Treehouse staff."

Sharon Raboin
Washington State Department of Labor & Industries

Treehouse Software, Inc.

Real World Series

(Part Seven of Several) by Joseph Brady and Daniel Sycalik

PENNSTATE 1855

This is the seventh installment in a continuing series of articles featuring **tRelational** and **Data Propagation System (DPS)**, Treehouse's ADABAS-to-RDBMS product implementation, in several "real world" environments.

tRelational autogenerates complete RDBMS schemata from existing ADABAS files and allows for easy mapping of ADABAS fields to already existing data warehouse or ERP schemata. After **tRelational** does the mapping, **DPS** can then materialize (initially load) and propagate (subsequently keep synchronized) the ADABAS data into the RDBMS without requiring direct access to ADABAS.

The following is a recent discussion between **Rusty Sodergren**, Lead Database Programmer Analyst at Penn State University, **Bill Cook**, Manager of Database Administration (DBA) for Enterprise Systems at Penn State University, and Treehouse Project Managers.

Please describe the Penn State University (PSU) organization.

Penn State is a multi-campus, public land-grant university that improves the lives of the people of Pennsylvania, the nation, and the world through integrated, high-quality programs in teaching, research, and service.

Please describe your department and charter.

Information Technology Services (ITS): ITS provides the infrastructure that enables Penn State students, faculty, and staff to make maximum use of the appropriate information technology tools in their learning, teaching, research,

"With the assistance of Treehouse's excellent remote and 'on site' support, we established a process that quickly was production ready."

(continued on page 4)

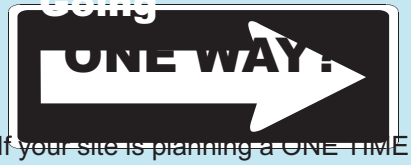
Editor's Notes by Joseph Brady

New Face at the Treehouse

We would like to welcome **Kevin Heimbaugh**, our new Senior Technical Representative.

In the past, Kevin has held important positions, both technical and non-technical, in the ISV business. His experience includes over 10 years with Software AG. He has been involved in sales, marketing, business development, product management, product support, pre-sales and post-sales activities, consulting, and even some DBA and application development work.

As a Senior Tech-Rep for TSI, Kevin will be involved in consulting engagements, pilot projects, sales calls, and product demos.



If your site is planning a ONE TIME migration off ADABAS to ORACLE, Microsoft SQL Server, Sybase, or IBM's DB2, **contact Treehouse Software today to find out about special pricing for tRelational** and **Data Propagation System (DPS)**, THE solution for migrating legacy ADABAS data into RDBMS-based Internet/Intranet/ERP/Data Warehouse applications.

Tree•times

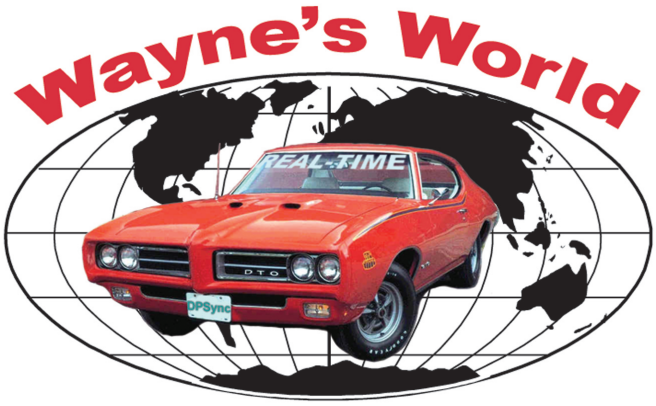
Editing, Writing, and Design
Joseph Brady

Contributors
Wayne Lashley, Chris Rudolph, Heather Snyder, Greg Such, and Dan Sycalik

Production and Distribution
Terri Hammerschmitt

Back issues available upon request. Documentation for all products is available in hard copy or on CD-ROM.

Hard Copy Circulation: 8,000



Wayne is Named the Director of Technical Operations at TSI

TSI technical activity is now headed by **Wayne Lashley**. Wayne has served as TSI's Group Leader of Consulting and Technical Sales for the past four years.

As Director of Technical Operations, Wayne now oversees product development, technical support and QA, marketing, documentation, pre and post sales consulting, and the Treehouse technical representatives. Wayne has more than 20 years of Software AG product experience in a 25-year IT career. Prior to joining Treehouse, Wayne held IT technical and management positions in academic, financial services, government, and primary industry sectors.

Wayne Presents the Latest TSI Products at the NATURAL Conference

Wayne Lashley, Director of Technical Operations, recently presented and demonstrated TSI's new offerings at the 12th NATURAL Conference in Cambridge, MA.

Wayne's presentation, ***Timely ADABAS Access with Real Products – Right Out of the Box***, addresses the challenge and urgency for integrating ADABAS-based data into modern IT architectures. XML applications, business intelligence systems, ERPs, and custom Java and .NET development all need timely access to the ADABAS databases that organizations depend upon. The presentation showed how Treehouse is addressing these challenges by offering products that do the job right out of the box with the comprehensive and robust features that you've come to expect from the leading ISV for ADABAS/NATURAL sites.

The first product presented was **DPSync**, which builds upon Treehouse's proven and highly-successful **tRelational/DPS** products by delivering near-real-time ADABAS-to-RDBMS data transfer. Much more than a mere change detection "solution", **DPSync** is a complete, mature, and fault-tolerant product that provides a metadata repository, ADABAS data content analysis, automated RDBMS schema design, ADABAS-to-RDBMS transformation mapping, native RDBMS bulk load (ETL), and periodic and near-real-time change data capture (CDC).

The second product covered in Wayne's presentation was **DPS X-Link**, the new ADABAS-to-XML tool from Treehouse. As enterprises begin to adopt XML-based technologies, those sites maintaining ADABAS systems are faced with extracting data and metadata and converting it into workable XML formats. To-date, home-grown and vendor-supplied "solutions" have led to a data/metadata code maintenance quagmire from which the XML revolution was originally engineered to free them. TSI's **DPS X-Link** product completely automates the process of converting legacy ADABAS data into XML documents and the corresponding PREDICT/ADABAS metadata into XML Schema. Authorized client applications can view an ADABAS datasource as one large XML repository with all database structures and metadata presented in XML Schema format instantly delivering ADABAS data in XML format.

Wayne's presentation from the 12th NATURAL Conference can be downloaded at: http://www.treehouse.com/proddownld.asp.

Boosting ADASTRIP's Power by Chris Rudolph

Treehouse is pleased to announce the availability of **ESTRIP**, the add-on enhancement to **ADASTRIP**.

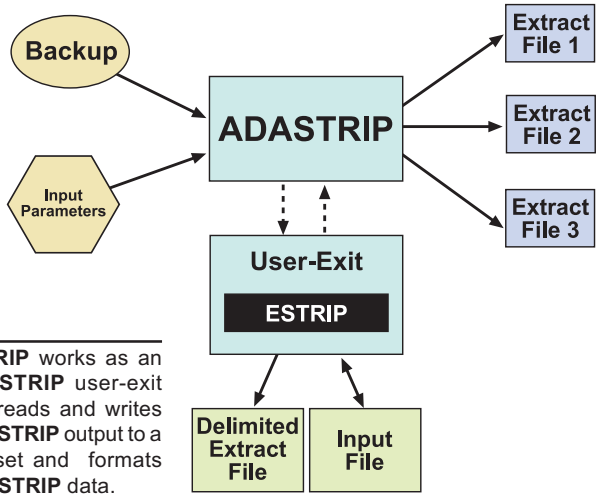
ADASTRIP has long been the DBA's preferred tool for extracting large amounts of data from ADABAS in a very short period of time. **ADASTRIP** users will immediately appreciate the features **ESTRIP** adds to the already powerful **ADASTRIP** product.

Using **ESTRIP**, one can strip data from an ADABAS file and then use that strip's output as input to strip corresponding records from a different ADABAS file. The input data can even be created from a non-ADABAS source, such as a SQL query from a data warehouse or an extract from a VSAM file. **ESTRIP** reads the input data and then uses **ADASTRIP** to extract the corresponding records from ADABAS.

The ESTRIP Advantage

Without **ESTRIP**, the user is faced with having to write COBOL or NATURAL programs to extract ADABAS data based on an input dataset. There is additional overhead in creating NATURAL DDMs, and in very little time, these additional DDMs can expand into a complex structure requiring constant maintenance. **ADASTRIP/ESTRIP** eliminates maintaining code or DDMs.

Many sites also have data warehouse projects in progress and need to quickly compare data from the warehouse against up-to-date data from ADABAS. With **ESTRIP**, it is easy to run a SQL query against a warehouse and extract data for comparison against ADABAS.



ESTRIP works as an **ADASTRIP** user-exit that reads and writes **ADASTRIP** output to a dataset and formats **ADASTRIP** data.

How does ESTRIP work?

Since **ESTRIP** acts as an **ADASTRIP** user-exit, **ESTRIP** enables the user to strip data from a file in which records are dependant on a previous strip or some external data source. This means that once a "parent" file has been stripped, the "child" files can be stripped so that related records are included in the output file.

ESTRIP will read the parameter cards that are either stored in memory for later use by the **ESTRIPX** exit or acted upon (i.e., reading, sorting, and creating the in memory tables).

For more information about **ESTRIP** or to arrange a free 30-day trial, please contact Treehouse Software.

A Lion in the Treehouse?

RedMane TECHNOLOGY

Treehouse has entered into a strategic partnership with RedMane Technology LLC, a Chicago-based professional services firm. RedMane's forte' is designing and building custom application solutions to help clients obtain competitive advantage while resolving their most compelling business issues. In addition to having a major practice area in ADABAS and NATURAL, RedMane has extensive experience in Web Enablement, Mobile Solutions, and Middleware utilizing Object-Oriented (OO) technologies.

By combining Treehouse competencies in data management/conversion and systems management products with RedMane expertise in application services, we provide valuable offerings for users of Software AG products. Benefits include:

- Investment protection of ADABAS and NATURAL systems.
- Innovative Web-enabled applications that are flexible, maintainable, and scalable.
- Robust Business Intelligence solutions for enhanced decision support.
- Comprehensive Replication/Migration Solutions that mitigates risk and speeds the deployment of newer, progressive technologies.
- Full-life-cycle Application Development, Application Integration, and Application Management solutions that reduce project life cycles and accelerate time-to-market.

To learn more, contact either **Greg Such** at Treehouse (gsuch@treehouse.com, 412-741-1677, x239) or **Brad Hagan** at RedMane (brad_hagan@redmane.com, 773-693-3919, x240).

outreach, administration, and support activities, as well as the cost-effective information technology resources required to support continuous improvement in the University's ability to fulfill its diverse mission.

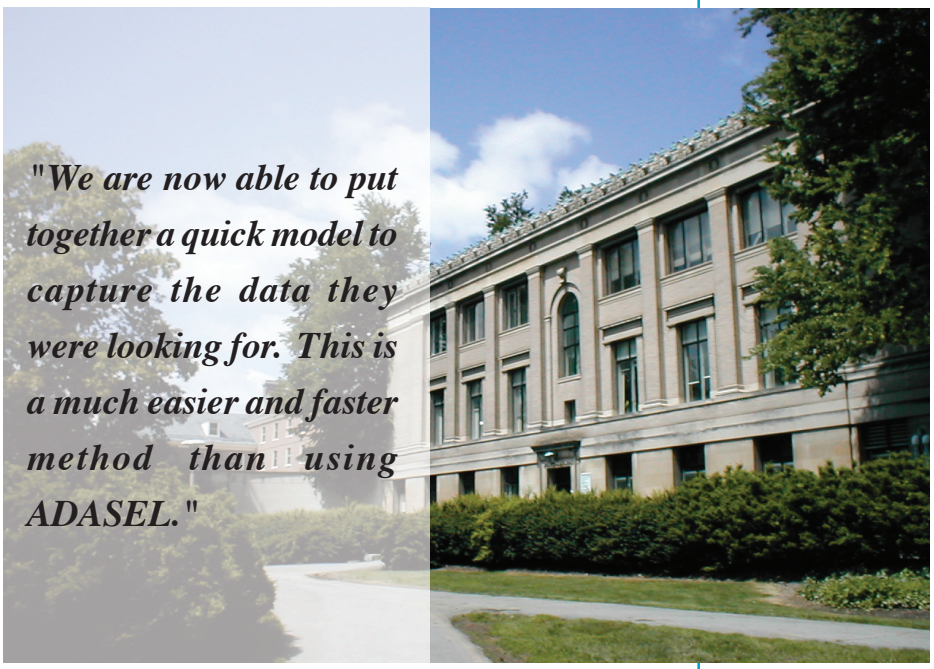
Please briefly describe your ADABAS/NATURAL environment.

We have two production databases containing the Integrated Business Information Systems (IBIS), the Integrated Student Information Systems (ISIS) and the Alumni Development Information Systems (ADIS).

Currently all of the core University business computing and information systems reside in ADABAS and have been programmed using NATURAL. The ISIS system includes student registration, enrollment, grades, scheduling, student loans, student aid, bursar, and admissions. The IBIS system includes accounting, budget, payroll, inventory, purchasing, human resources, financial systems, and electronic forms. The Alumni system contains all kinds of alumni and development information, donations, giving, etc.

What prompted PSU to evaluate tRelational and DPS?

PSU has been considering various tools over the years to enhance propagation of ADABAS data to other environments, including various RDBMSs and flat files. Like other shops, we have also suffered from the processing overload of NATURAL extracts as a means of data extraction and propagation.



We spoke to **Mitch Doricich** (Treehouse National Sales Manager) and company during a Conference in 2002 about **tRelational** and **DPS**. Previously, at the 2001 NATURAL Conference, we attended a **tRe/DPS** demo conducted by **Wayne Lashley** (Treehouse Director of Technical Operations) and **Brian Johnson** (from Cutler-Hammer).

We talked to other vendors as well. The Treehouse solution had the largest market share, and we knew it was a seasoned product.

Can you describe the tRe/DPS Pilot experience?

The pilot was very successful. We learned a lot about the product set and its capabilities. With the assistance of Treehouse's excellent remote and 'on site' support, we established a process that quickly was production ready.

We completed an end-to-end nightly production propagation process that executed in 15 minutes, which included the FTP and the load to ORACLE. This was a significant time-savings compared to the existing methods consisting of a NATURAL extract that was then used to perform a full refresh of the data.

We were very pleased, and your technical support always had an answer for every question. We were impressed with the data transformations and the features of the product, and we were also impressed with the excellent support.

Would you give us some details on the PSU tRe/DPS Projects?

The first project was for the Undergraduate Admissions Office at Penn State. They had developed an application that generates 'match codes' for names and addresses of all prospective students and applicants for Admission. These match codes are used to identify redundant information in the database. For instance, a person could be in the database as "J. Doe" and "John Doe" or as "Bill" and "William." Or, they might have two different addresses in the database, but they both are for the same person. The match codes identify these redundancies and help prevent PSU from corresponding with the same person as if they were two or more different people. This reduces the number of duplicate letters, emails, personal contacts, etc., thus saving time, effort, and funds and helping PSU present a more professional image to students, prospects, and their families. It also prevents important information from getting placed on two different records, such as SAT Scores getting assigned to "William" and Applicant Information getting assigned to "Bill."

(continued on page 5)

Real World Series *(continued from page 4)*

The Undergraduate Admissions Office, in conjunction with ITS Enterprise Systems DBA group, created a **tRelational** data model to capture all of the names and addresses from the "person" file and "correspondence" file. Initially, a **DPS** materialization of the data was done to capture all of the names and addresses (filtering on home address only). Then, daily propagation started, which captured all name and address changes. Both the materialization and propagation processes use ACTALODF, which creates a fixed-length flat file of the information to be passed to a NATURAL program that subsequently reads the file and generates match codes. Additionally, the propagation process uses APC (ADABAS PLOG Consolidation) to capture only the final image of the record, rather than all of the changes that were done in a day, since it is not necessary to know about each change, only the final result.

It sounds like ACTALODF and APC are going to be well-used features *(see notes on ACTALODF and APC on page 6)***.**

Yes. Both ACTALODF and APC are providing a great benefit and are now being used for additional processes, and plans are being made to expand the use of these features. In fact, we recently modified the original data model to feed some new applications that are related to a very important, high profile project at Penn State. PSU is in the midst of converting the primary faculty/staff/student identifier from the Social Security Number to a new Penn State ID Number. The original data model developed for Admissions was modified to capture some additional name/address data. An application was then written to parse this data to identify name, address, and SSN changes, and then update two new ADABAS files with the SSN changes and match code information.

There are a couple of other projects currently being discussed that will also use ACTALODF/APC as input to current processes. One use will be to populate report datasets (in SAS) that the Office of Student Aid uses as their data warehouse. They currently use NATURAL extracts and the "pull of data" is a full refresh done weekly. Student Aid plans to implement **DPS** to capture changes daily so their information is more up to date. This simple process enables the Office of Student Aid to avoid doing a full weekly refresh of the data, saving CPU cycles, as well as providing much better service to their office.

Finally, we have already benefited greatly by using ACTALODF/APC for some internal troubleshooting. Periodically, programmers come to the DBA with a request to identify information contained on a particular PLOG. We are now able to put together a quick model to capture the data they were looking for. This is a much easier and faster method than using ADASEL.

We'd also like to mention that Treehouse wrote a special column routine called COLAMAX for us to capture the last occurrence of a PE, which was very important to this project and other future projects.

Is tRe/DPS helping Penn State in any other areas?

Yes. In addition to the projects previously mentioned, Admissions requested more frequent updates of their data that currently resides in the Penn State Data Warehouse. In particular, they were looking for daily updates to the data, as well as expanded access. A number of methods for meeting these goals were discussed, including modifying the current NATURAL extract process, possibly using ADASEL, or using a product called SAS/Access. Our DBA staff was approached by Admissions for some help using SAS/Access, and we mentioned the **tRe/DPS** trial to them. It was decided that we'd try some test models using the **tRelational** Autogen feature. Admissions was thrilled with the results.

"...we began daily propagation of inserts, updates, and deletes. The daily process is very fast and has been virtually trouble free."

As a result of Admissions' request, we began the process of creating normalized staging tables for their office. We did an Autogen of nine ADABAS files. The whole process was very simple and straightforward. The most time-consuming part of the project is running the file analysis, which is required for Autogen, but once that is finished, the rest is gravy. **tRelational** does everything, including creating the DDL and GENning the parameters used as input by **DPS**.

(continued on page 6)

Real World Series *(continued from page 5)*

We materialized the nine ADABAS files into 162 MS SQL Server tables. Then, we began daily propagation of inserts, updates, and deletes. The daily process is very fast and has been virtually trouble free. Admissions has written a number of VIEWS of the normalized data that in turn are being used to produce daily Admissions reports both at our University Park and Commonwealth Campuses.

Next, we in DBA will be working closely with Penn State's Data Warehouse Consultant in selecting warehouse tables that may benefit from population from the staging tables. In addition, we will work together in analyzing the feasibility and benefits of adding new tables to the warehouse or restructuring selected existing tables. We anticipate that this will be able to be done in a more timely and efficient manner by extracting data from the staging tables.

Ultimately, we do not plan on having direct access available to the staging tables, though they may be made available for some ad-hoc reports by a limited number of power users. The long term plan is to expand the staging tables to include more Autogenned ADABAS files that may then be used to populate additional existing data warehouse tables or use them to create new tables.

What were the major decision points or factors in purchasing tRe/DPS?

The major point, and it's really fundamental, is it allowed us to accomplish things that we could not do otherwise. For instance, we could not populate daily updates to the Admissions data on the Data Warehouse due to the time it takes to run the NATURAL Extracts. We had investigated alternative solutions without success.

How would you rate the products features, functionality, and performance?

We would give them an A grade. We really have not yet come across a requirement that we could not find a way to accomplish.

How would you rate the Treehouse support, skills, and knowledge related to the processing requirements?

We would give an A+. As good as it could possibly be. We deal with several fine vendors and you guys are at the top. The timeliness, the quality, and the knowledge are excellent. The development of the new COLAMAX ETR met an important requirement, which was very helpful, and we really appreciate this level of customer support. In particular, **Dan Sycalik**'s support and knowledge has been key to our success.

Do you have future plans for additional tRelational and DPS projects?

Absolutely! As a matter of fact, there are a number of projects in the works. Number one is the replacement of the current process of populating the transcript table in the Data Warehouse. The other is converting the NATURAL extract process for the Office of Student Aid Data Warehouse. Recently, we met to discuss an upgrade from weekly to daily updates.

We are beginning to get the word out through various means, including presentations to committees (University-wide Data Access Committee) and internal presentations (Enterprise Systems Group).

Would you recommend tRE/DPS for other clients considering ADABAS data transfer?

Most definitely!

Notes on ACTALODx and APC

by Daniel Sycalik

ACTALODF delivers a fixed format full row image from **DPS** Propagation (Change Data Capture) processing, as opposed to the standard SQL DML (Inserts, Updates, Deletes and commits). Additionally, a UID_INDICATOR column may be mapped to provide a "U"pdate, "I"nsert, or "D"elete flag to indicate the activity. The full row image may be utilized for change reporting, application system interface, interface with other third party software, or programmatically manipulated for additional processing requirements.

Alternatively, ACTALODD generates a column delimited full row image (e.g., CSV formatted). Several clients have utilized ACTALODx routines to customize their **DPS** Implementation. In fact, one client justified the purchase of **tRelational** and **DPS** to replace a daily change reporting process that took over six hours to execute (note: **DPS** executed in 10 minutes).

The ADABAS PLOG Consolidation (APC) utility consolidates transaction activity against the same ADABAS record to condense the **DPS** Propagation output. APC intelligently factors mapping criteria, such as Primary Key and Foreign Key constraints, so APC may not always condense down to a single DML statement or ACTALOD row image. However, APC has often resulted in an average of 25 – 30% consolidation. This reduction of transaction activity yields a significant time and work reduction during the RDBMS processing.

These product features operate right "out of the box" and deliver functionality that many might consider "out of the box" thinking for ADABAS to RDBMS processing.

tRelational has capabilities that you might not be aware of that take it beyond ADABAS-to-RDBMS modeling and mapping. **tRelational is also a data analysis tool for ADABAS** that can help you with performance issues, file structure, and data integrity.

tRelational Provides Robust Data Analysis

tRelational is principally a mainframe-based NATURAL application providing a range of analysis, modeling, and mapping functions in both online and batch modes. Installation of the basic mainframe **tRelational** facilities enables commencement of the process of identification of ADABAS data sources and resolution of issues pertaining to them.

Resolving Discrepancies During File Implementation

tRelational file implementation retrieves file structure information from PREDICT data definition entries (DDEs) and the ADABAS Field Definition Table (FDT) and creates an "implemented file" entry in the TRE-DICT part of the **tRelational** repository. Discrepancies between PREDICT and the FDT are flagged in the Implemented Field Summary and can be investigated and resolved in order to ensure accurate representation of the data structures.

File Analysis

The results of **tRelational** file analysis are important not only to guide RDBMS modeling and mapping, but to uncover potential data quality issues. Such issues may already be causing adverse effects in your operational applications now, and can have disastrous effects in ADABAS-to-RDBMS projects if they are not identified and resolved early.

By default, file analysis processes all records in physical sequence. However, subsets of records may be analyzed by specifying a descriptor (and optionally a descriptor value range). This is particularly useful when multiple logical "record types" exist within a physical ADABAS file.

The following table illustrates the value of file analysis:

Analysis	*Relevant statistics*	*Primary purpose(s)*	*Potential data quality issues that may be identified*
Repeating Field	• High occurrence count • Average occurrence count • Count of records with at least one occurrence • Count of records with multiple occurrences	• Support normalization vs. denormalization modeling decisions • Estimate number of rows in normalized “child tables”	• Unused MU/PE fields • MU/PE fields that are populated differently from expectations
Alphanumeric	• Non-null value occurrence count • Highest length found • Average length	• Support RDBMS datatype (CHAR vs. VARCHAR or other) and length selection in modeling	• Unused fields • Fields where maximum length exceeds defined length
Descriptor/ Superdescriptor	• Total values found • Distinct values found	• Support selection of primary key for RDBMS table(s)	• Duplicate values in descriptors and superdescriptors assumed to be unique • Records lacking a value in a required descriptor or superdescriptor component

Tree•times

Treehouse Software Products

ADABAS-to-RDBMS Data Transfer:

DPS - ADABAS-to-RDBMS data materialization (ETL), replication, and propagation (CDC) software

DPS X-LINK - Instant XML-based access to ADABAS

DPSync - Near-real-time ADABAS-to-RDBMS data propagation (CDC) software product set

tRelational - ADABAS modeling, mapping, and data analysis tool; **DPS** parameter generator

tRelationalPC - Windows-based graphical interface to make the tasks of modeling and mapping even simpler

Treehouse Remote Access (TRA) - Middleware that allows **tRelationalPC** to communicate with **tRelational** on the mainframe.

Data Integration:

iBahn - Integration suite that connects data within the enterprise or between business partners

UNIX:

SEDIT - XEDIT and ISPF/PDF compatible editor for UNIX and Windows

S/REXX - REXX-compatible language for UNIX and Windows

S/REXX Debugger - Optional graphical debugger for **S/REXX** programs

Software AG Related:

ADAREORG - File reorganization tool for ADABAS

ADASTRIP - Data extraction utility for ADABAS

AUDITRE - Generalized ADABAS auditing facility

AUTOLOADER - ADABAS file automatic unload/reload/dump utility

CHART for NATURAL - NATURAL application analysis and documentation tool

N2O - NATURAL application change management system

N2O/3GL - 3GL support within **N2O** for PANVALET, LIBRARIAN, ENDEVOR, and PDSs

PROFILER for NATURAL - NATURAL quality assurance and testing tool

SECURITRE - ADABAS and NATURAL security interface to RACF, ACF2, and TOP SECRET

TRIM - ADABAS and NATURAL performance monitor

Phone: (412) 741•1677 **Fax:** (412) 741•7245 **E-mail:** tsi@treehouse.com **Web:** http://www.treehouse.com

TREEHOUSE SOFTWARE, INC.
409 Broad Street, Suite 140
Sewickley, PA 15143 USA