

NatQuery

The

Data Extraction Solution For ADABAS™

Overview.....	2
General Features.....	2
Integration to Natural / ADABAS.....	6
NatQuery Modes.....	7
Administrator Mode.....	7
FTP Information.....	8
Environment Configuration Information.....	8
End-User Mode.....	9
Demo Mode.....	10
NatQuery for End-User Extraction.....	10
NatQuery for Data Warehousing.....	11
Data Extraction.....	11
Change Data Capture.....	11
Integration with Extraction, Transformation and Loading (ETL) Tools.....	11
General Information.....	11
Workstation Requirements.....	12
Server Platform Requirements.....	12
Software AG™ Product Dependencies.....	13
Supported Source Environments.....	13
Security.....	13

NatQuery Abstract

Overview

NatQuery is a workstation-based, Graphical User Interface (GUI) application that has been specifically designed to support data extract processing against data stored in Software AG's ADABAS database, or any other data source that can be accessed using Software AG's Natural programming language. Originally designed as a "user-friendly" replacement for Software AG's product Super Natural, NatQuery is applicable for End-User extraction, Data Warehousing, basic Web Publishing, and basic data replication. NatQuery is applicable to use on any platform that also supports ADABAS & Natural with options for scheduling, reporting, and a variety of data retrieval options. Leveraging ADABAS and/or Natural utilities as well as the Natural Programming language assures compatibility with ADABAS as well as an inherent familiarity of our tools by your Administrators. NatQuery supports data extraction from ADABAS in two ways:

- The dynamic generation of data extraction programs written in Natural™, Software AG's 4th generational programming language,
- The dynamic generation of parameter cards for the execution of specific ADABAS utilities as well as the generation of Natural programs to properly process the output of these utility executions.

From its inception, NatQuery has been designed to fully automate and simplify data extraction from ADABAS. It is a tool designed to be:

- Easily used by an End-User with minimal training
- An indispensable ADABAS administrative tool
- Easily installed
- Easily administered and maintained
- Easily integrated into existing desktop data-processing applications
- Robust and performance oriented.

General Features

Please note that Software AG has developed ADABAS / Natural to run on many different platforms. These platforms will call for certain "general" features of NatQuery that may or may not apply to your particular environment. If you have questions regarding available features, please contact NatWorks.

- **Flexible Extraction using straight Natural or the use of ADABAS utilities and Natural**
Ability to extract data from ADABAS through the use of 100% generated Natural programs, or through the use of specific ADABAS utilities that are supported by 100% generated Natural programs.

NatQuery Abstract

- **“Online” or “Offline” Extraction from ADABAS**
When ADABAS utilities are used for extraction, NatQuery supports this in either “offline” or “online” modes. In “offline” mode, NatQuery supports extraction from ADABAS by utilizing ADASAV (ADABAS Backup) tapes as input. In “online” mode, NatQuery supports extraction through straight generated Natural, or through the use of ADABAS utilities such as ADAULD and ADACMP with the output of these utilities then being processed with generated Natural.
- **Performance Sensitive Natural Generation**
When Natural is used for extraction, NatQuery generates performance-sensitive Natural version 2.x structured programs with the ability to automatically generate the various appropriate Natural data access statements using descriptors, super-descriptors, hyper-descriptors, sub-descriptors, ISNs and phonetic-descriptors as access keys.
- **Multi-File Extraction**
In conjunction with utilities or using straight Natural, NatQuery has the ability to automatically generate complete and ready-to-execute Natural data extraction programs that support the logical access of up to 5 related “primary” files with each of these 5 files having the ability to further access any number of "lookup" references (i.e. code-to-text translations), within a single generated program.
- **Straight-Forward Integration to the Natural / ADABAS Server**
Automatic or manually controlled integration between the user's workstation and the remote server through the use of basic FTP services, allowing for both the upload of NatQuery requests, the downloading of request output, or FTP to a different platform.
- **Scheduled Query Submission**
NatQuery allows the user to set up scheduled submission of their queries with a wide variety of parameters: Once Only at a specific time, Daily in specific intervals, Weekly by specific time and day(s), Monthly by specific time and date(s).
- **Automatic Leveraging of Performance Feature**
Automatic support for the Multi-Fetch feature of ADABAS, a feature that significantly reduces execution time and CPU overhead when Read Physicals are required.
- **Integration to PREDICT**
Automated integration to PREDICT, with the ability to make PREDICT information available within the NatQuery environment as well as the ability to provide and pass PREDICT information into Data Warehousing products.
- **Automatic Execution of Extraction Requests**
Auto execution of extract requests in batch allows full control over potential impact to production online systems and ensures the best performing Natural environment.

NatQuery Abstract

- **Automatic Handling of required Job Control Language (JCL) / Scripts**
Automatic generation of Job Control Language / Scripts to execute NatQuery processes, based on a template-based approach. Automatic calculation of the total logical record length (LRECL) of extracted data and estimation of the number of records returned, allowing for efficient and appropriate use of disk space on the server platform.
- **Natively Handles ADABAS Data Types**
Full handling of all ADABAS data formats including Date and Timestamp with translation into user-friendly strings.
- **Full Support for Recurring Fields**
Full support for "recurring" data structures including Periodic-Group fields (PEs), Multiple-Valued fields (MUs), as well as MUs in PEs. When output, these structures may be "concatenated" or "flattened" in the resulting data extract at the user's discretion.
- **Automatic Data Conversions**
Full and automatic handling of data conversions from internal ADABAS formats in EBCDIC to ASCII equivalents such as Date, Time, and Packed to Numeric (including automatic handling of sign bytes).
- **Support for Variables**
Full Support for user-definable "variables" including; Constants, Expressions, Redefinition or Dynamic (system-date based and user-supplied).
- **Automatic Update Capable**
Automatic update capability that allows distributed NatQuery End-User client workstations to have new versions applied from a single point of administration.
- **Sharing Extraction Definitions**
Full ability to store query specifications, generated programs, and full JCL execution streams on user-specified paths, allowing for the reusability and later modification of entered query specifications as well as the sharing of extract specifications.
- **Support for Delimited or non-Delimited Extracts**
Support for both non-delimited and delimited data extracts. Delimiters are user specifiable, thereby creating extracts that support a wide range of available workstation data handling tools, including Seagate's Crystal Reports, MS Access, and MS Excel.
- **Automatic Integration to EXCEL**
Direct integration to Microsoft Excel, with automatic creation of appropriate column "headers".

NatQuery Abstract

- **Automatic Integration to ACCESS**
Direct integration to Microsoft Access, with the ability to automatically create an Access table, append to an existing table, or merge into an existing table.
- **Delivery of Data that is “Load Ready”**
NatQuery can deliver data directly to a Data Mart or Data Warehouse platform in a format that can be immediately processed by a RDBMS Load utility. For some RDBMS extracts such as Oracle and SQL Server, NatQuery has the ability to automatically load this data.
- **Integration to ETL Tools**
Direct Integration to IBM’s WebSphere Data Integration Suite - formerly Ascential Software’s DataStage Suite, Pervasive’s Business Integration suite, or any other Data Warehouse product that can accept COBOL File Definitions (CFDs).
- **Data Replication / Data Update Capability**
Used in conjunction with the NatCDC product, NatQuery can provide basic transaction replication from one ADABAS to another, and can additionally process sequential input files with optional Store, Update or Delete processing against a given ADABAS file.
- **ADABAS Data into XML / XSL**
NatQuery has the ability to convert, in a single JCL stream or Script, ADABAS data directly into XML on the server. Alternatively, NatQuery further supports the conversion of ADABAS data into XML with an accompanying XSL template when directed to do so on the user’s workstation such that the extracted data is immediately ready to be published to the internet.
- **Ability to Process Sequential Files**
NatQuery has the ability to create mainframe based sequential extracts that can then be immediately used as a data source for NatQuery. This allows for extracts which are based upon a sequential “driver file”, with the full ability to link the sequential file to other ADABAS files to accomplish expanded extracts. Additionally, NatQuery can be instructed to read virtually any flat sequential file whether created by NatQuery to accomplish the same tasks.
- **Data Discovery Functions**
Data Discovery functions allow an End-User to analyze ADABAS Repeating Fields (Multi-Valued Fields and Periodic Groups) and allow for a thorough understanding of how a given ADABAS application uses these types of fields. Another function allows for Alphanumeric Field Lengths to be analyzed for maximum length used, number of null fields and other pertinent information. These functions augment existing capabilities of NatQuery that analyze the use of ADABAS “key” fields such as Descriptors and Super-Descriptors.

NatQuery Abstract

Integration to Natural / ADABAS

NatQuery on a PC workstation integrates to the remote Natural / ADABAS server platform using automated File Transfer Protocol (FTP) over TCP/IP or automated File Copy across a network. Physical executions of all NatQuery requests occur in “batch”. Within this document we shall assume that FTP will be used as the preferred integration process.

To support the batch execution of extracts, NatQuery utilizes generic JCL / Script templates that are specific to a given task. When NatQuery is instructed to process a request, NatQuery generates the required objects and / or commands, and then inserts these into an appropriate JCL / Script template. This template is then parsed for pre-defined dynamic field tags, with these dynamic field tags allowing for generic JCL / Script templates to become specific to a given user / user request. This processing results in fully customized and executable JCL / Script. How this JCL / Script stream is then submitted to the platform’s internal reader is based upon an administratively set option.

If an organization allows for the direct submission of JCL into the remote platform’s internal reader queue (for MVS this would be JES, for VSE this would be POWER), then NatQuery can be configured to FTP the executable JCL *directly* into a reader queue where it can execute. If the Natural server platform is UNIX / Linux or Windows, then automated execution can be accomplished via one of several Remote Execution approaches; RSH, REXEC, or local poll and submit process executed at intervals. Please note: NatQuery easily supports manual integration / execution approaches as well.

When a specific user’s request executes, the requested data is written to a user-specific server dataset, and the request’s execution additionally updates a server-based “log” file that is specific to the requesting user with information concerning the executed request. This server-based user log file is retrievable to the user via FTP, and when this is done NatQuery will compare the contents of the server-based log to the contents of a workstation-based log (that is updated as requests are sent for execution). This comparison provides the user with the relative status of each request. When a request is shown as completed, the output of the request can be retrieved by the user via automated FTP.

Once the requested data is retrieved by the user to the NatQuery workstation, further automatic processing may occur if previously specified, such as the loading and presentation of data in Excel, or the loading of data into Access.

NatQuery Abstract

NatQuery Modes

There are three “modes” that NatQuery may operate under:

- Administrator
- End-User
- Demo

The mode that NatQuery will operate under is designated by the use (or lack of use) of a “License Key”.

For the purpose of general demonstration, NatQuery can be executed without any License Key, in which case NatQuery will operate as a Demo. When provided with an appropriate key, NatQuery will operate in either Administrator mode or End-User mode.

License Keys are available from NatWorks, Inc. Depending upon the requirements / conditions of your NatQuery installation, at a minimum a License Key will be provided to support an Administrator installation(s), and an additional License Key to support End-User installations.

For questions concerning the availability of License Keys, please contact NatWorks, Inc through our website located at URL <<http://www.natworks-inc.com>>.

NOTE 1:

Depending upon the requirements / conditions of any given NatQuery installation, NatWorks may provide you with License Keys that are time-expiring. In these cases, NatQuery will be fully operational under the appropriate mode until the pre-set expiration date. After this date, NatQuery will automatically revert to operating in Demo mode.

NOTE 2:

It is possible to switch the mode under which any given installation of NatQuery is operating by providing the installation with an appropriate License Key. In this way time-expired installations can be re-activated, or an Administrator installation may be switched to an End-User version or vice-versa.

NOTE 3:

For Data Warehousing applications of NatQuery, only a single Administrator version of NatQuery is usually required. End-User versions are designed to enable End-User extraction of ADABAS data, and may not be required for this use.

Administrator Mode

The Administrator Mode is designed for use by a designated administrator, and it serves to build and manipulate categories of information that directly influence NatQuery and its interaction with the remote Natural / ADABAS database. Additionally, the Administrator Mode has the ability to execute / move requests against the ADABAS environment and therefore contains all functionality found in the End-User mode of NatQuery.

NatQuery Abstract

There are two general categories of information that a NatQuery Administrator mode is designed to capture:

- FTP Information
- Environment Configuration Information

FTP Information

The FTP Information section will supply NatQuery with the information required to connect to the remote ADABAS platform via FTP. This includes FTP-related information concerning the server platform, the designated users of NatQuery, and the JCL/Script templates that will be executed on the server platform.

Environment Configuration Information

The Environment Configuration Information section pertains to information that is directly related to the source ADABAS database(s). This includes the capture of:

- **Data Definition Modules (DDMs)**
DDMs are a basic building block for Natural access to ADABAS. DDMs are used by NatQuery to gain information on specific ADABAS files. DDMs can be obtained from the ADABAS platform automatically from within NatQuery, or DDMs can be imported into NatQuery.
- **File Relationships Information**
File Relationship Information is optional, but if provided, gives NatQuery the ability to allow the automatic logical linking (or “joining”) of multiple ADABAS files together into a single data extract request. The information when captured describes how any given DDM relates to other DDM(s), the access paths (keys / indices) that support the logical linking, and how these access paths are initialized. With this information available to NatQuery, an End-User is completely shielded from having to deal with these complex issues.
- **Descriptor Statistics**
Descriptor Statistic information describes all of the access paths (keys / indices) available into a given file and describes the I/O characteristics of the access path. This information can either be automatically generated or may be manually supplied. With this information available, NatQuery gains application-specific intelligence, and can determine, based on user-supplied Selection Logic, what the optimal access path is to resolve a given data extract request.
- **Occurrence Information**
Occurrence Information describes the default occurrence specifications for recurring fields, as well as their specific maximum limits. This information then controls how a given data extract request will obtain and handle the recurring fields structures in ADABAS such as Multi-Valued fields (MUs), Periodic-Groups (PEs) and even MUs in PEs. With this information available, an End-User is

NatQuery Abstract

completely shielded from the need for in-depth knowledge of these complex structures.

- **I/O Parameters**

I/O Parameters generally describes program generation “rules-of-thumb” for when generating Natural I/O statements. These parameters will influence when one access method will be used over another.

With the above information captured, NatQuery is provided with “application intelligence”, an intelligence that can be equated to that of a skilled Natural programmer who understands the complexity of a given database application file structure. This intelligence is then utilized by NatQuery to enable End-User generation of optimized single and multi-file data extract programs on demand, in a manner that completely shields the end-user from the complexities of how source data is physically stored, and the mechanisms needed/used to extract the required data.

End-User Mode

The End-User version is designed specifically to allow “End-Users” to build extraction requests against the remote ADABAS database. It cannot be used to perform any required Administration tasks.

By using the Environment Configuration that has been built using an Administrator version, the process that an End-User must perform to specify any extraction request is reduced to its most basic elements:

- What file(s) does the desired data reside in
- What data fields need to be retrieved
- What additional variable fields should be defined (this is optional)
- What selection logic should be applied to retrieve the specified data
- What the target of the data extract should be

The end-user component supplies the graphical means of capturing the above information, with this information being referred to as a request or Query Specification. Query Specifications are handled in a typical Windows fashion, with the standard functions of Open, Save, Save As, Delete, etc.

When handling recurring fields, these fields are automatically presented to the end-user in terms of their administratively controlled “occurrence” defaults, which greatly simplifies the understanding of these types of fields on the part of the end-user. At the end-user’s discretion, these occurrence default values can be easily over-ridden (within the administratively controlled maximums), but with default values provided, users do not need an in-depth knowledge of these structures.

Once a Query specification is entered, and prior to the actual generation of a Natural extract program that satisfies the query specification, the query can be designated by the user to be a specific type of extract program. The currently supported extract types are:

NatQuery Abstract

- Data Download to PC File
- Data Download into Microsoft Excel
- DWH Software Extract
- SourcePoint™ Extract program
- Data Download to Microsoft Access
- RDBMS Loading
- Download to XML File
- Any2XML Extract
- Extract to XML on the Server

Demo Mode

The Demo version is designed specifically to allow prospective customers to appreciate the look and feel of NatQuery and how it operates. It must be noted that the Demo version's use is restricted in the following ways:

- The Demo version does not allow for the creation of a customer-specific environment. While all administration functions can be seen, these functions cannot be utilized to create a customer-specific environment configuration.
- The Demo version is limited to generating Natural data extraction programs against an installation provided "demo" environment configuration, a configuration that uses the Software AG example DDMs of EMPLOYEES, VEHICLES and SAG-TOURS.
- The demo version cannot be used to communicate to a remote Natural / ADABAS platform.

NatQuery for End-User Extraction

The concept behind the design of NatQuery was simple:

- Reduce, to an absolute minimum, the knowledge that an End-User would require to successfully extract data from ADABAS

This design approach was a departure from other available ADABAS-specific data extraction tools; tools that had proven themselves to require significant amounts of education on the part of the End-User to use, let alone use effectively.

NatWorks Inc feels that we have delivered on this basic design concept through the implementation of an Administrator component. This component allows an Administrator to provide NatQuery with knowledge that is specific to any given ADABAS application. Once captured, NatQuery can utilize this knowledge to guide, facilitate and radically simplify the specification of a given data extraction request.

NatQuery Abstract

NatQuery for Data Warehousing

One of the most difficult tasks with any Data Warehousing effort is the implementation and management of the physical data capture processes. There are two major aspects in considering these processes:

- Data Extraction (for initial data loading / full data refreshing)
- Change Data Capture (CDC)

Data Extraction

For general Data Extraction, NatQuery is all that is required. For maximum flexibility, NatQuery allows for the generation of Natural data extraction programs. For maximum performance, NatQuery supports the generation of the parameters required for specific ADABAS utilities, as well as the generation of Natural programs that process the data that these utilities deliver.

Change Data Capture

For Change Data Capture, NatWorks offers the product NatCDC, an add-on product to NatQuery. Using the combination of NatQuery and NatCDC, an organization can easily access transactional or the “changed data” of ADABAS – data stored in ADABAS Protection Logs (PLOGS). For further information on NatCDC, please contact NatWorks or refer to the NatWorks website located at URL <http://www.natworks-inc.com>.

Integration with Extraction, Transformation and Loading (ETL) Tools

In many cases, the delivery of extracted data is only one step in a total Data Warehousing effort. Once the data is available, there are often requirements to further process this data through transformations and / or scrubbing, as well as the processing required to physically load this data onto the target platform.

To facilitate the further processing of NatQuery-provided data into ETL tools, NatQuery has the ability to generate “interface files”, or files that describe the physical layout and content of NatQuery-produced data extractions. NatQuery supports several interface file types, such as COBOL File Definition files (CFD) and SQL.

Once NatQuery has generated CFD or SQL files, these files can then be imported into any one of a number of ETL tools. Once imported, these ETL tools gain the capability to immediately process the extracted data in the required manner.

General Information

NatQuery is written in Visual Basic, and is delivered as a 32-bit application that has been designed to execute on any Microsoft Windows 32-bit operating system.

NatQuery Abstract

Workstation Requirements

The following are the suggested minimum requirements for workstation installation:

- 486 processor (Pentium recommended)
- 6MB Hard Disk free (minimum)
- 16MB RAM (64MB or higher recommended)

To achieve automatic integration to a remote Natural / ADABAS platform, NatQuery utilizes File Transfer Protocol (FTP). This means that in order to achieve this integration, the workstation must be physically connectable to the server environment over TCP/IP, and have appropriate hardware / software installed on the workstation that will support an FTP client.

The physical implementation of FTP with NatQuery is achieved through program calls to a Microsoft-supplied Dynamic Link Library (DLL) called wininet.dll. Specific versions of this DLL are required depending upon the host Natural / ADABAS operating system. The following requirements apply:

- For VSE and UNIX servers, wininet.dll must be version 4.0 or greater
- For MVS servers, wininet.dll must be version 5.0 or greater

Microsoft generally supplies Wininet.dll as part of the Windows operating system; NatQuery does not supply this DLL. If this DLL is not on your PC, or an older version is installed, you may obtain this DLL by installing the latest version of Microsoft's Internet Explorer product. Internet Explorer is obtainable free-of-charge through Microsoft's website.

Server Platform Requirements

NatQuery achieves integration to a remote Natural server platform through the use of File Transfer Protocol (FTP). This means that the server platform must support FTP Server connections, along with associated hardware / software.

NatQuery can utilize FTP connectivity in one of two ways:

- By performing FTP operations directly into "batch" on the server's operating platform. For example, directly into JES on MVS systems or into POWER for VSE systems. To accomplish this, the server platform must be configured to accommodate these batch requests – in some cases this may be disallowed for security concerns.
- By performing FTP operations into pre-defined disk files on the server's operating platform which are then subsequently picked up and submitted to the server platform through a batch-based NatQuery server process. This server process is provided with a NatQuery installation, and may be run intermittently, continuously, or at pre-scheduled intervals to allow requests to

NatQuery Abstract

be processed.

For VSE systems, it is currently a requirement that VSAM be available to contain required server-side files.

Software AG™ Product Dependencies

The following table summarizes the associated Software AG products and their relative dependencies on the server platform.

Product	Features/Comments
Natural	Required – version 2.x or higher
ADABAS	Optional – version 6 or higher – 3.1 or higher for UNIX/Windows
PREDICT	Optional
Natural Security	Optional
ADABAS Security	Optional
Natural Interface modules	Optional - only required for non-ADABAS data sources

There are no Software AG product dependencies for NatQuery on the workstation.

Supported Source Environments

The following table summarizes the operational environments that support Natural and are therefore suitable for use with NatQuery.

Platform	Operating System
IBM	MVS (OS/390), VSE, AS400
DEC	Open VMS
UNIX	UNIX – where Natural is also supported
Siemens	BS2000
Fujitsu	MVS derivative
Hitachi	MVS derivative

There are no Software AG product dependencies for NatQuery on the workstation.

Security

If security is an issue at a customer site, this is most likely already being addressed by using a security package such as RACF or Top Secret, and additionally perhaps through the use of Natural Security and / or the security features of ADABAS. Regardless of the security measures employed, the physical execution of any NatQuery request must ultimately occur under a known system User-ID and corresponding password. NatQuery therefore fully embraces the security measures that are already in place on the server system.