Building a Data Warehouse

with LANSA

A guide to successful
business user
data access

LANSA

A *advanced software made simple*

Building a Data Warehouse with LANSA

First edition, March 1996.

Please note: While every attempt has been made to ensure the accuracy of the information contained in this white paper, readers are advised to draw their own conclusions from multiple references.

Building a Data Warehouse with LANSA

A guide to successful data access

In today's fast moving world many business users want to access AS/400 corporate data. This guide examines the key success factors in implementing business user data access and shows how LANSA can rapidly build a practical Data Warehouse.

It covers the following topics:

- **Too much data, not enough time or information**
  IT has traditionally created reports for business users but the continually rising demand has driven the search for a means for business users to create their own reports.

- **End user query - only part of the answer**
  Data query tools promised business users the ability to extract information to support decisions from the mass of stored data. Implementation has often been less than successful. While data access tools have been easy to use, the corporate file structure has often been too complex for users to understand. The IT department often offloaded a reporting backlog but took on a support headache.

- **End user query - only part of IT's role**
  Data access by business users is important but, uncontrolled, it can cause security and performance problems for production systems. IT must balance the competing needs of conflicting groups while sharing limited resources.
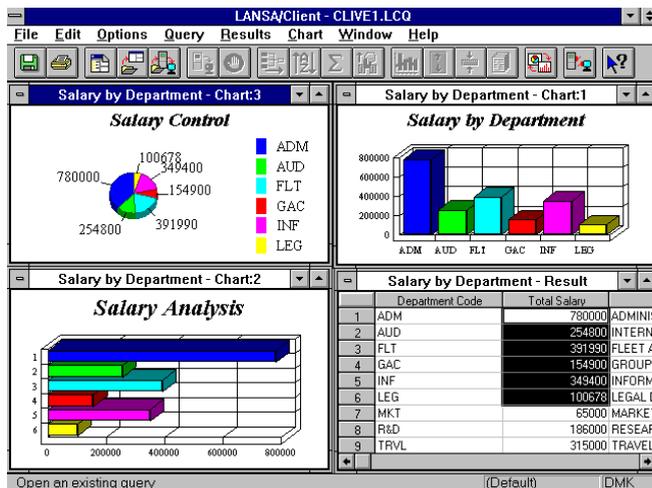
An easy-to-use query tool is only part of a complete data access solution

- **Effective data access requires a Data Warehouse**
  Databases designed for operational systems follow "normal form" theory. These databases are often inappropriate for business user access in view, complexity and performance terms.

  A Data Warehouse transforms and groups the corporate data into business terms that users understand. IT also gains better control over system performance and secures protection of production files.

- **Selecting a Data Warehouse strategy for your AS/400**
  All sites can benefit from some form of Data Warehouse. Many AS/400 sites may only need to define simplified files but others may need to do much more to achieve effective data access by business users.

- **Key steps to data access success with LANSA**
  Follow these guidelines to achieve successful data access by business users within your organization.

- **The LANSA data access family of products**
  LANSA is ideal for creating a Data Warehouse. LANSA's query, report and charting tool is the easiest to learn and use AS/400 data access tool on the market.

- **LANSA services**
  Professional services can be the key to a successful implementation. A comprehensive range of LANSA services is available.

Operational data

Data transformation
by LANSA

Departmental Data Marts
created by LANSA

Customer summaries

Product summaries

Monthly summaries

Information access
•LANSA/Client
•LANSA/Server

LANSA Repository
Data Warehouse
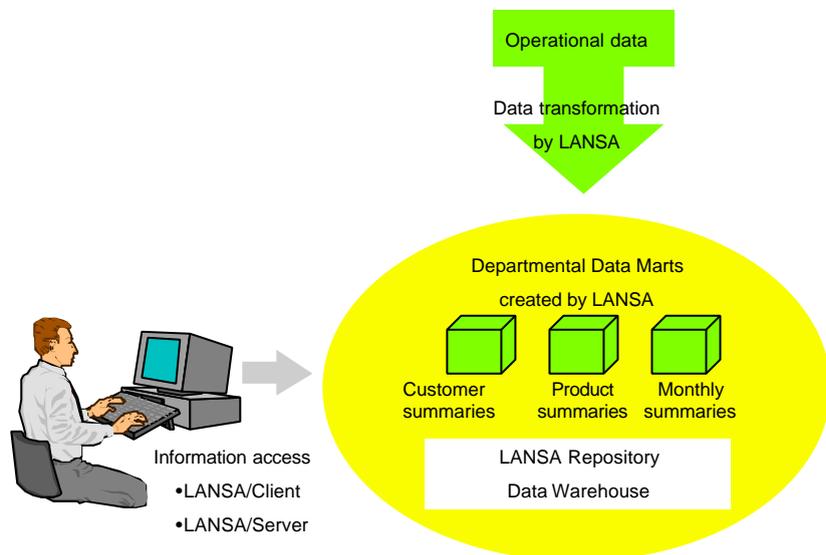
LANSA offers a complete data access solution, creating an effective Data Warehouse and providing an easy-to-use data access tool

Too much data, not enough time

Traditionally, the Information Technology department automated business processes. As business becomes more competitive, processes are becoming more flexible and less clearly defined. Now, IT only has time to define and automate the core business processes, ensuring data integrity, security, backup and recovery disciplines are in place.

Business users often have increased decision-making authority and fulfil more varied and ever-changing job roles. As a result, users must adapt IT systems in order to rapidly respond to ever-changing business pressures.

If successful, sharing the automation workload between the IT department and business users delivers a more adaptable solution in a shorter timeframe. Business professionals can change the delivered solution as their needs change. IT is freed from the reporting backlog to concentrate on the development of mission critical applications.

Forrester Research reported in January, 1995 that as much as 70% of the application backlog consists of building queries and reports.

30%

Reports 70%

Grown from 60% in 1994

Source: Forrester

January, 1995
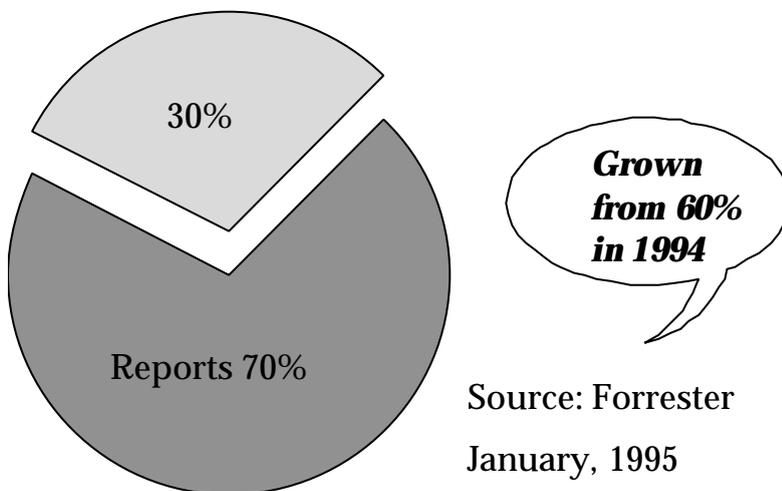
Business user data access is needed to reduce the reporting proportion of the IT application backlog

NOTE : Report is a generic term for information presented on GUI or paper which may contain text or graphics.

End user query - only part of the answer

There are many challenges to solve before business users can easily access corporate data and create their own reports. Simply providing business users with an "easy-to-use" data access tool may only replace an IT reporting workload with a support workload.

As a business user attempts to access corporate data, support questions arise:

- users have difficulty understanding application files
- users have difficulty analyzing complex data
- users have difficulty with inconsistent and incomplete data
- users have problems with the supporting systems environment

The need for easy-to-understand fields and files

- **"I want to look at customers - what is the GL301L1 file and is GLCSTID the customer?"**
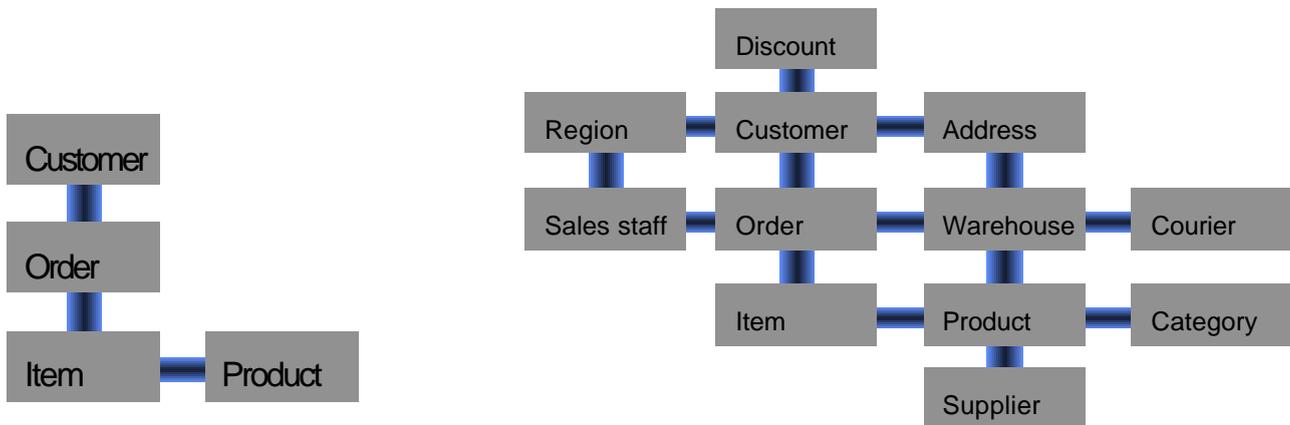
The typical demo database is simpler than a real database

Business users often reject the complexity of actual systems (even without multiple currencies or other complications)

- **Complicated file structures and too many files**
  Customer data may be spread across seven (sometimes many more) files because commercial transaction processing systems are designed with normalized files. Normalized files make sure that field updates and deletes alter only the relevant data and that insertions don't create redundant data. This forces data to be spread across many files. Programmers are familiar with normalized files but business users are often overwhelmed by the complexity of operational file structures.

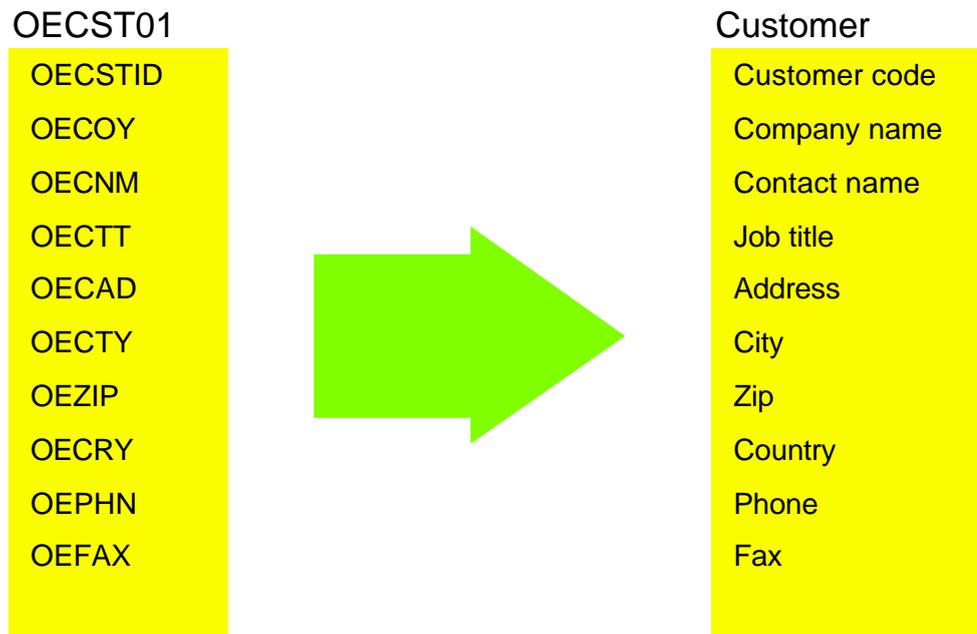 The view seen by business users can be simplified into more natural files such as a customer file, a product file and so on. Since data access involves only data retrieval the data does not need to be normalized. Normalized files are only desired if data is updated, deleted or inserted.

  Business users need education in the meaning of the files and fields that compose their business system. Otherwise, IT will receive frequent support calls.

- **Indecipherable file and field names**
  Programmers define file and field names according to site standards, for example, GL301L1 may be a file within the General Ledger system.

OECST01	Customer
OECSTID	Customer code
OECOY	Company name
OECNM	Contact name
OECTT	Job title
OECAD	Address
OECTY	City
OEZIP	Zip
OECRY	Country
OEPHN	Phone
OEFAX	Fax

For business users, LANSA's Repository can be used to provide business-friendly names to avoid viewing the programmer naming standards

Business users expect to see familiar business names such as customer file. Not all tools have a Repository or Catalog but one is necessary to rename fields and files to familiar terms.

- **No Help text**
  Even if the file view has been simplified and familiar business terms have been used to rename the fields, help text is needed to allow users to better understand the corporate data definitions. For instances, does "Sales" mean "Invoiced goods and services" or "Paid invoices". It is an important on-site complement to formal business system training. And it must be up-to-date or users will rapidly lose confidence in it.

The need for easy data analysis

- **"I don't understand SQL - how can I analyze the data?"**

SQL allows programmers to query databases but most business users won't understand it.

Modern query tools allow business users to easily create queries without knowing anything about SQL. Ease-of-use should not restrict query power. Files will probably need to be connected to find, for example, how many products a customer has ordered during June.

The data access tool should understand the links between files so that the user need only mouse click and follow easy prompts to drill down through multiple files or to seamlessly join related data from different files. Reporting and charting the retrieved data should be just as easy.

Ease-of-learning is heightened if the data access tool follows the familiar Windows-style interface, particularly if a spreadsheet is used to manipulate data. Most business users are already very familiar with spreadsheets. If they can copy the retrieved data from the data access tool into familiar tools learning is further minimized.

Good data access tools also come with extensive on-line reference or tutorial material to further guide users.
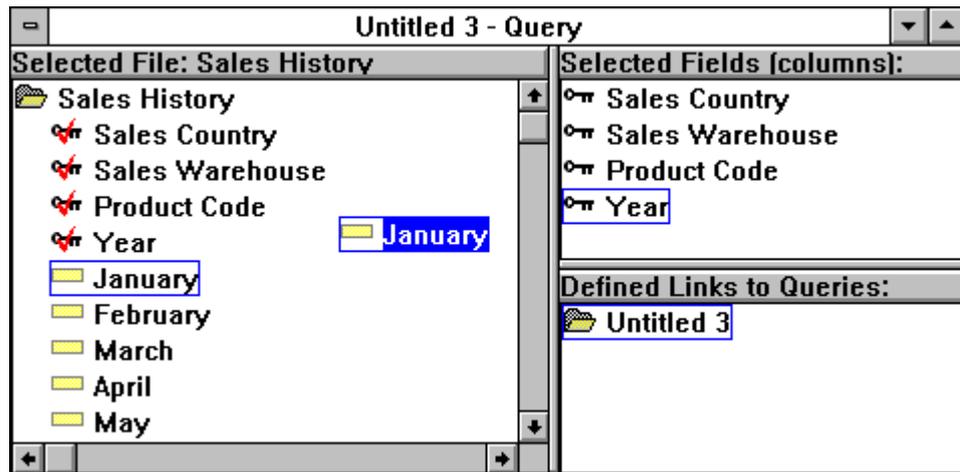
With LANSA/Client, a business user can query with point-and-click ease, including automatically linking across multiple files.

The need for consistent and complete data

- **"Why does a field have different values in different files?"**

  That can't happen, can it? Well, unfortunately it can. Business systems evolve over time and are written by different programmers and it is a rare computer that does not contain conflicting data in separate business systems.

  Data can be inconsistent between packages acquired from different vendors and data stored with similar names may have quite different uses in different files as used by different departments. For example, the Sales department has a vested reason to make "sales" look as high as possible but the Distribution department wants to make sure "sales" reflects everything that is shipped - whether it is sold by a salesperson or not.

  Another source of business user confusion occurs if records stay in the database even though they are "deleted". Often there is a "Marked for deletion" field to signify deletion. Auditors can find all records - deleted or not. Business users may create invalid summaries and reach damaging conclusions. Either educate users about their files or remove confusing data from the files they access.

  Many real systems actually contain a plethora of "flags" meaningless to business users. These are added either to save disk space or, pragmatically, to add features to an application after its initial implementation.

  A data field may also contain an embedded data array of values. For example, the "Budget values" field may actually contain twelve values, one for each month. The beginning of "Budget values" may even be an internally described field that defines what is in the rest of the field according to codes determined by the programmer. Embedded data arrays and internally described fields are created by programmers to save disk space and to improve database performance. They are relatively common.

  LANSA's Repository can be used to decode fields such as "Budget values" into separate fields for each month but other data access tools will return only one combined, confusing field to the user.

- **"I'm using ODBC- why can't I find my data?"**
  Using ODBC to access the AS/400 may not be a complete answer because significant AS/400 business data is often hidden from ODBC.

  AS/400 data inaccessible to ODBC
  Non-AS/400 or generic ODBC data access tools cannot access AS/400 non-relational data:

  - Files that have originated from a System/36 consist of internally defined fields (physical fields composed of multiple logical fields) that are unknown to ODBC. An ODBC driver will only retrieve a string of data as a single physical field. It can't know how to separate the contents into distinct fields. The user must decode the definitions manually or acquire a tool with the capability of decoding the definitions from a specification of the internal format.

    The LANSA Repository can decode System/36 data format.

  - AS/400 multi-member files can only be accessed by examining each member as a separate file. By default SQL/ODBC accesses only the first member. It can't access other members.

    The LANSA Repository can treat each member as a separate file and access all members of a multi-member file.

  - AS/400 multi-record format files contain different logical formats in different records, usually identified by a code at the beginning of the record. The codes must be interpreted before the different formats can be accessed as distinct logical views.

    The LANSA Repository can treat each format as a separate logical file.

  Data on other platforms inaccessible to ODBC
  Many AS/400 sites also have other Server systems such as MVS or UNIX storing non-ODBC data. As an integrated relational database machine the AS/400 is ideal as a central site to integrate data from other platforms.

  IBM offers a range of products to migrate data stored in various formats to the AS/400.

ODBC data field names are indecipherable programmer terms
The AS/400 has a Data Dictionary for translating the terms that programmers define into terms familiar to business users. (These definitions are usually called metadata - data about data in Data Warehouse literature.) ODBC can access this dictionary. However, the implementation of a dictionary varies between database systems. Data access tools that can only use an ODBC data source all force the dictionary to be separately defined on the PC. This creates change control problems because the PC catalog(s) for all users will need to be kept in step with changes made on the AS/400.

With LANSA, the business user receives Windows point-and-click productivity and consistent business-friendly field and file definitions while IT obtains the systems management ease of the AS/400. LANSA's client/server Repository allows central definition and maintenance with an AS/400 set of skills on the AS/400. At the start of each data access session using LANSA/Client, the appropriate portions of the Repository are transferred to the client. All access to the client Repository that requires access to the database includes a version comparison so that an updated Repository is only loaded when the copy on the client is no longer current.

- **"How can I answer complex nested questions?"**
  Exploring business opportunities may mean the user needs **ALL** of the customer sales data not only the data from the Sales system but also customer data kept in various other systems. They expect an integrated and complete view of all the data related to the customer. They simply don't expect the systems to be as complicated as they often are.

  Business users often find that the answer to a question raises many other questions and a data access tool must be able to dynamically regroup data by different dimensions such as customer groupings, product groupings, by time period of year, month etc.

  To follow analyses to a conclusion requires both summary-level data and detailed data. One way is by using a tool that can dynamically create summaries from detailed records. Another way is by defining summary files as part of the system. A pragmatic mix is often the best solution. Defining too many summary files can use too much disk space and users always seem to want a new summary that was not specified when the system was designed.

Country	City	Product Description	Year	January
Argentina	Buenos Aires	Cryptograph Procesor		278750.00
		Data compression		303520.00
		Optical Data Server		418000.00
		LAN Connector		186200.00
		Colour Ink Jet Print		103200.00
		Printer		205440.00
	La Plata	Cryptograph Procesor		537230.00
		Remote Control Unit		1067640.00
		Workstat'n Controler		271920.00
		DASD 320		337440.00

Dynamically regrouping data is easy with LANSA/Client's dynamic creation of multi-dimensional views

The need to train users about Network access

- **"I often can't connect to the AS/400 - why?"**

  Business users normally understand Windows applications quite well. Network access is often expected to be just as easy as using any Windows application. It isn't.

  Reliable and repeatable connection to the AS/400 can only be straightforward if the unintegrated network communication components are carefully installed and configured by knowledgeable staff. Network communications can be complex and IT should consider specialized consulting services to recommend proven standards and implement an effective installation, training and support program.

  It is tempting to try to avoid connection issues by using a file transfer utility to download files and then directly access the data from the PC. This may be a good alternative for smaller data extracts but it is rarely a complete solution.

  Corporate data usually resides on the AS/400 so that IT can keep it up-to-date with ongoing business transactions. A file transfer creates multi-platform update problems that complicate the disciplined controls needed if business decisions are to accurately reflect the current data status.

  The PC is a completely different architecture to the AS/400. For example, the data format on the PC is ASCII format where as AS/400 data is EBCDIC format. The AS/400 packed data type does not even exist on the PC. Data access security and system availability are also more controllable when data resides on an AS/400.

  LANSA automatically shields you from the complexity of platform differences, allowing smooth access to AS/400 data. However, the PC environment is not as integrated and friendly as the AS/400 and training in underlying network complexities is essential.

End user query - only part of IT's role

"How can I keep all users happy without compromising security and performance?"

IT must provide security and performance for all users. Business user data access will reduce IT workload only if user requests for on-going help are less than the work required to create and maintain IT-created reports.

Share support resources across all users

An easy-to-use data access tool helps reduce requests for on-going help. Train your users in the tool and in the underlying application and the meaning of its fields and files. Train your users about your PC and Network environment.

Many sites have extensive libraries of pre-defined Query/400 queries. A PC-based query tool that can run Query/400 queries may at first seem attractive, but its SQL-like language will probably create a higher support load. It is probably a better approach to let users continue to run the Query/400 queries using a 5250 emulator and give them a modern PC data access tool that can be used to more easily define a similar query while adding greatly enhanced graphical reports and charts.

IT must juggle limited resources across many users

Many business users are too busy managing and selling to find the time to learn a data access tool, no matter how easy to use. These users still require IT to create reports. This workload can be reduced by using a comprehensive 4GL such as LANSA to create and maintain reports more rapidly. Also, LANSA's exception processing can raise alerts as email messages to warn of, for example, low stock conditions.

Ensure security for all users

Many data access tools break the security offered by software packages because the tools bypass application menus to access files directly. Historically, application menus restrict user access. Although the AS/400 offers file-level security it is not normally used because it often conflicts with menu security.

LANSA offers security that permits authorized PC users to access their files without conflicting with menu security or forcing you to define system-wide AS/400 security.

Balance performance for all users

Business users expect rapid response time but this must be balanced against the needs of competing system users. Data access can impose demands on DB2 for OS/400 (or any relational database) that impact production transaction systems. If a business user is trying to join two multi-thousand record files together at the same time as an important order is trying to update one of the files, significant transactional delays can occur. This can be minimized or avoided in several ways.

Firstly, some data access tools use AS/400 native data access facilities that are more efficient than non-native SQL or ODBC.

Secondly, a data access tool should execute queries in a separate sub-system from batch or interactive jobs. Such a separate sub-system can be configured to minimize impact on other users.

Thirdly, a copy of the detailed transactional files can be created either on the same AS/400 (to avoid file contention) or on a separate AS/400 (to avoid central processor or DASD contention).

Note that it is not recommended to copy the files to a PC unless you are confident you have chosen a PC database and PC hardware configuration that can comfortably contain the large volume of corporate AS/400 data. Significant PC management skills may be needed to tune the PC database and maintain backups and coordinate regular data propagation from the AS/400 to the PC.

A complete Data Warehouse solution not only simplifies files for business users but also addresses performance and security issues.

Effective data access requires a Data Warehouse

"OK, effective data access requires more than a query tool but what is a Data Warehouse and do I really need one?"

An effective Data Warehouse is a complete solution, delivering effective data access for business users. The purpose of a Data Warehouse is:

- **to provide an easy-to-understand view of corporate data**
  The most important objective is that the intended business users must be able to understand, query and analyze the data.

- **to provide a consistent data view**
  Data should be "cleaned" (inconsistencies eliminated) to assist with accurate analysis and to help make the data easy-to-understand.

- **to provide a complete view**
  A complete view allows complex questions to be answered but too much data can increase complexity. To avoid complexity, the requirements of the intended department of business users should be understood and specifically targeted, but care must be taken not to include unnecessary data.

These objectives are designed to improve the productivity of business users. IT has additional objectives:

- to provide data access without impacting the security or performance of other users.

- to provide data access within given resource constraints.

To understand how a Data Warehouse can meet these objectives we must first look at what a Data Warehouse is.

"What are Data Warehouses, Data Marts and OLAP?"

The modern view of a **Data Warehouse** is a read-only database designed for business user data analysis. It presents files that reflect the structure of the business such as customers and products and contains a complete and consistent set of data.

Historically, Data Warehouses began at mainframe sites to address the need to integrate data from many different databases. Some data may have been stored in hierarchical databases such as IMS, other data in DB2 relational tables, still more data in VSAM, ISAM or even flat file formats. Much of this data was stored across multiple platforms. The original purpose of a Data Warehouse was to get a complete set of data in one place. Data was copied into a common format and definitions cleaned for data consistency. These early Data Warehouses often took too long to create and rapid business change rendered many obsolete. Their creation often lacked business focus.

A **Data Mart** is the term coined to describe a smaller, more pragmatic Data Warehouse. A Data Mart contains the data needed by one department, say Marketing. It doesn't need to contain all the corporate data, but only the data Marketing staff require such as customer sales information. It may not need production, distribution or billing data. It usually has a strong focus on achieving business benefits. Data not required to achieve the identified business goal is not included in the Data Mart.

Online Analytical Processing (OLAP) refers to the data analysis business users perform and the business-user-friendly databases they need. The term is a deliberate contrast with Online Transaction Processing (OLTP), the term commonly used to describe day-to-day commercial processing and its associated highly normalized databases.

"What is an EIS, DSS - do I need more?"

Modern data access tools are more flexible, offering improved ease-of-use and more analytical power.

An **Executive Information System (EIS**) originally meant a set of pre-defined summary files and graphical business reports. First, being pre-defined, they often raised as many questions as they answered. Secondly, as management hierarchies flattened many more staff needed business reports. They became Enterprise Information Systems but gradually gave way to more flexible data access tools. Thirdly, it is generally necessary to analyse using detailed data as well as summary data, not because business users will want to look at individual detailed records, but because the data must often be summarised very precisely. If a user wants to see how many sales of a product happened by day following a specific media promotion detailed data will be required.
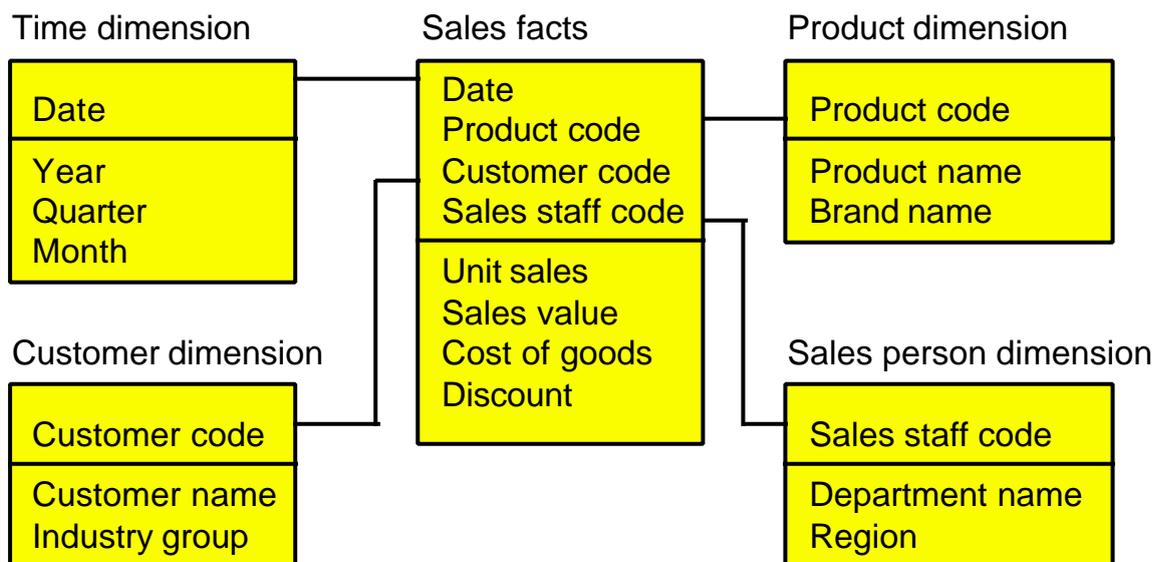
Decision Support Systems (DSS) offered more analysis capabilities and were used by more technical staff responsible for detailed decision-making. Early tools often required users to know SQL or other techniques with equally steep learning curves. Again, they have gradually given way to more flexible and easy-to-use data access tools.

"What does an OLAP database design look like?"

A Data Warehouse database uses an OLAP design. An OLAP or multi-dimensional design is very different from a normalized design and commonly uses a **"star join schema".** A central fact table contains detailed data and dimensional tables index it. One dimensional table exists for each of the business users natural groupings. For example, common dimensions are Time (year, quarter, month summaries), Customer (Type summaries),Market (District, Region summaries), Products(Category summaries). This approach makes it much easier for a business user to simply select, say customer, and choose the facts about customer they want to query instead of trying to navigate across many tables to bring all the data together.

A multi-dimensional database can be created as permanent files on the AS/400 using the integrated relational database to create a "star join schema" or a flexible multi-dimensional database can be dynamically created by LANSA's data access tool, LANSA/Client on the PC.

Note that, although often repesented as a cube, a multi-dimensional database usually has more than the three dimensions of a cube. It has as many as is natural and desired by its business users.

Time dimension

Date
Year Quarter Month

Sales facts

Date Product code Customer code Sales staff code
Unit sales Sales value Cost of goods Discount

Product dimension

Product code
Product name Brand name

Customer dimension

Customer code
Customer name Industry group

Sales person dimension

Sales staff code
Department name Region

A multi-dimensional design, shown as a "star join schema", reflects the structure of the business and is much easier for business users to understand. Files are denormalized into a central fact table and dimension tables index it. Dimensions are chosen based on the user's preferred natural groupings.

Evolving a Data Warehouse for your AS/400

1. **All AS/400 sites already have a form of Data Warehouse**
   As an integrated relational database machine the AS/400 has data stored in a consistent format already. However, the view exposed to business users needs simplification because the fields and files follow programmer naming conventions and files are normalized.

   LANSA's Repository can be used to assign business-friendly names to fields and files. These definitions are then used by LANSA's data access tool, LANSA/Client to make data access easier for business users.

2. **Integrate data from non-AS/400 systems**
   Multi-platform sites will want to take advantage of the AS/400 architecture to move data from other systems. Recent enhancements to the AS/400 provide parallel query processing and symmetric multiple processing to satisfy the largest of Data Warehouse requirements.

3. **Simplify the user view of data**
   A modern data access tool like LANSA/Client can simplify file and field names. Performance and security issues can be addressed by LANSA's Repository which can provide a security layer while renaming fields and files to business terms and adding explanatory Help text. Users will need to be educated in the application structure.

   LANSA can also simplify the user view of files by defining automatic joins for easy file navigation and by defining predetermined join fields that display fields from multiple files as one simplified file. For example, customer name and address is usually stored in a customer file, separate from the order file which contains order details and the customer code. Pre-determined joins allow a business user to view order details and customer name and address as if they come from a single file even though they are actually stored in multiple files. Business users gain simplified access while IT optimizes both performance (pre-determined joins use native RPG access and are much faster than SQL or ODBC) and disk space (by avoiding the performance contention between data access and production systems that is otherwise solved by having them access separate copies of files).

   Some data access tools such as LANSA/Client allow grouping and summarizing to dynamically create summary views of data. This facility can be used to prototype the dimensions needed for a Data Warehouse. Stable requirements can be converted into AS/400 files using LANSA. This has the added benefit of improving data access performance because summaries are already stored in the file instead of being calculated by the business user's data access tool at query time. The trade-off is increased disk space to store summaries. LANSA's predetermined joins can also be used to create summary fields without increasing disk space.

4. **Control performance and security by dedicating a copy of the operational files for data access**
   Create a copy of operational files and allow business users access only to the copy. Performance and security is easier to control (particularly if copied to a separate AS/400) but this step does not by itself simplify the file complexity. It is a natural additional step if you find performance issues remain after implementing a data access tool directly against operational files.

5. **Start small by defining and creating a departmental Data Warehouse (called a Data Mart)**
   A Data Mart is a small Data Warehouse targeted only at a single department's specific requirements. It should contain both detailed transactional data and summary files required by your highest priority business area which may be, for example, the marketing department's view of customer. For most sites the dimensions such as Product groupings or Customer groupings will be available from existing systems such as the Sales History system.

   LANSA's predetermined join fields can transform the normalised History system files into a user view that reflects the business structure (a Data Warehouse's multidimensional database). This method optimizes disk space but means that summary fields are calculated each time. Alternatively, define the multidimensional database using LANSA and populate it with LANSA triggers from the operational files. As a trigger can be run as a delayed batch job, performance impact is minimized.

   Data should be cleansed for consistency across multiple packages as the Data Mart is created. It can be difficult to reach agreement because different meanings often exist between different departments. For example, sales may mean the net of revenue less returns to the Finance manager. Sales to the Distribution department is what needs to be delivered. Sales to the Sales department is the amount committed to by clients. Restricting the Data Mart to a single department helps avoid these interdepartmental conflicts. LANSA can cleanse the data as the Data Mart is created.

   A solution that may suit some but not all sites is to create a temporary Data Mart - create the summary files infrequently, say quarterly (choose a frequency that allows meaningful trends but not so frequent that effective disk space is increased), produce trend reports and then delete the Data Mart. Marketing can analyze trends without permanently consuming disk space.

6. **Define and create an enterprise Data Warehouse by progressively implementing Data Marts**
   This is a serious undertaking similar to creating an enterprise data model for operational systems. The benefits can be high but the analysis time may mean results are not available quickly enough. Avoid this by carefully expanding the Data Marts in priority business order until the enterprise view is as complete as the business benefits justify. Functional overlap is likely and should be anticipated if merging is to be successful. Also, each Data Mart will continually change its data storage requirements as the underlying business evolves.

Selecting a Data Warehouse strategy

All business users can benefit from the ease-of-use aim of a Data Warehouse. The extent the ideas are adopted depend on:

- **the business benefits to be gained from data analysis**
  Greater business benefits justify more resources to create a Data Warehouse. More volatile, highly competitive industries are more likely to yield higher business benefits.

- **the technical skills of the users**
  Less technical users require greater simplification of the corporate data. IT users may not need any simplification of the data.

- **the complexity of the data to be analyzed**
  More complex data will need more cleansing and simplification. For example, data may need to be converted from different data formats such as from mainframe data bases or VSAM files into AS/400 relational data. Data may also need to be transformed from different logical formats, for example, different geographic regions often have different logical data definitions.

Most **stand-alone AS/400 sites** will probably only need to implement file simplification and only for a few departments where business benefits can be clearly identified.

Sites with **multiple AS/400s** or **multiple heterogenous platforms** may need to integrate data onto one AS/400 to obtain a complete data view.

"Must I buy a separate AS/400?"

LANSA's Repository can simplify the user view of data into a Data Warehouse's multi-dimensional database, often without compromising the performance of the operational systems. LANSA's predetermined joins use fast native access but mean business users view data as if from a single file even though the data may be stored across multiple normalised files. LANSA's data access tool, LANSA/Client runs in a separate subsystem to allow better control over the relative performance of data access users.

Sites with a higher number of data access business users and a higher rate of usage may want to consider a separate AS/400 for the Data Warehouse, especially if their production system is close to capacity. As the cost of computing has come down, data access can now be economically offloaded to a separate system, making both data access and production performance much more controllable.

Key steps to success

7. Gain executive sponsorship with realistic expectations

An executive sponsor is important to the success of a Data Warehouse because benefits are often intangible and difficult to quantify. In organisations where business departments control their share of the computing budget, approval will be easier to obtain because the benefits tend to be self-evident to business managers. Even in this case it is necessary to give executive sponsors a realistic expectation. Otherwise expectations can be impossibly high to meet.

Key steps:
1. Gain support
2. Identify benefit and priority by department
3. Select complete set of tools
4. Analyze requirements
5. Prototype
6. Train business users

Data Warehouses do not give users all the information they need (because there is always some information not collected). Sufficient detail and sufficient breadth is required to answer the questions posed by the sponsoring executive. If a question asked by an executive requires uncollected data additional IT resource will be required but the executive may well perceive lengthy delays at a crucial business time. This may be perceived as project failure unless expectations are carefully managed.

Professional services should be considered to achieve the user-oriented analysis of the executive's Data Warehouse requirements.

8. Identify and prioritize departments according to the business benefits of data access

Examine your backlog of report requests and you will probably agree with many sites that it is primarily the finance and marketing departments that demand ad hoc reports.

Finance personnel revise cash flow projections based on operational data and need to access AS/400 data to copy it into their spreadsheets for further analysis. The benefit of data access may well be better cash flow as overdue receivables are identified more easily.

Marketing personnel today have dynamic job responsibilities as they try to delight customers with high levels of service. They have two kinds of typical questions:

- **Ad-hoc questions**: for example, where is the missing product delivery?

- **Discovering business opportunities**: for example, can a comparison of product sales across geographic regions uncover new sales opportunities?

The benefit of data access can be improved customer relations and increased sales.

Categorize the report backlog and try to quantify and prioritize the business benefit of data access.

9. **Choose a complete solution - not simply a data access tool**
A data access tool provides Windows ease-of-use but only a complete LANSA data access solution simplifies the files, allowing true business user productivity and helps IT control both security and performance.

10. **Analyze and evolve your requirement for a Data Warehouse**
Consider the complexity of your applications and the technical skills of your business users and decide how much data cleansing and file simplification you require. Data summaries may not be needed except on an ad hoc basis. If so, they can be calculated by the data access tool. However, pre-calculated summaries stored in the Data Warehouse can improve performance dramatically at the expense of extra disk space. For example, it is commonly required to compare current month sales with the average for the previous year. Calculating the average may require summarizing very large numbers of detailed records unless frequently required summaries are stored in the Data Warehouse. A separate AS/400 may be considered to provide more control over performance and security.

11. **Prototype your data access solution as a departmental Data Mart**
Choose the highest priority business department, prototype both the nature of data access by creating a Data Mart and remember to create and validate the training and support procedures. Refine both as new business areas are implemented. Creating a departmental Data Mart also helps avoid interdepartmental conflict over "correct" data definitions.

12. **Train business users well**

- **Application file structure** - particularly if you choose not to simplify the files. This is essential but is frequently overlooked.

- **PC and Network Communications** - Create standards, teach and be ready with support. Consider professional services to train your users and correctly install your data access tool and supporting environment.

- **Data access tool**- Every tool requires some training for optimally effective use. Training is required initially but also continually for infrequently used options and newly released features. Review the supporting material such as on-line tutorials or references and the education workshops available.

LANSA/AD - fast and simple Data Warehouse creation

LANSA/AD contains a powerful 4GL for building an AS/400 Data Warehouse. In 1995, the leading newspaper *Midrange Systems* voted LANSA the AS/400 4GL of the year.

- **Transform non-relational AS/400 data** - LANSA is particularly well suited to working with existing databases whether initially created or not by LANSA. As well as normalized database files (managed by DB2 for OS/400), LANSA can work with virtually all native AS/400 files, including System/36 files, multi-member files and multi-record format files. Such files can be easily transformed into relational files that can then be viewed either by LANSA/Client or by any query tool that can access DB2 for OS/400 files such as the many ODBC data access tools.

- **Transform non-AS/400 data** - LANSA can either access non-AS/400 data directly using ODBC or be used to cleanse the data after propagation to the AS/400 with IBM products such as DataPropagator.

- **Simplify files** - define business-friendly terms, automatic file navigation, pre-determined joins and Help text that can be reused by both business users and programmers alike. Business users gain the benefits when using LANSA's data access tool, LANSA/Client and programmers can reuse the same definitions to create customised programs and reports.

- **Powerful triggers** extend native AS/400 capabilities to allow updates to operational files to propagate to the Data Warehouse using a delayed batch job. This avoids any impact on production systems while keeping the Data Warehouse current.

- **Centrally defined Repository-maintained business rules** cleanse data to ensure Data Warehouse consistency.

- **Alerts** or exception reporting can be directed to email to eliminate the need to manually analyze many reports.

LANSA/Client - turning data into information
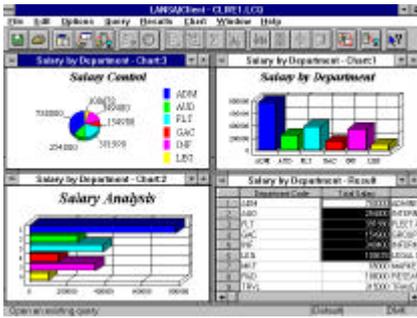
LANSA/Client is an easy-to-use query, reporting and charting tool for accessing AS/400 data.

- **Familiar business terms** - Repository-defined field and file descriptions present terms familiar to business users.

- **Report templates** - users can run standard reports or readily view and tailor reports.

- **File simplification** - users view simple consolidated (denormalized) files prepared by LANSA/AD as the Data Warehouse or use predetermined file joins to view many files as one.

- **File filtering** - users can reduce the number of files to be viewed with file filtering.

- **Automatic joins** - users can simply select fields from multiple files to create a consolidated report.

- **Drill downs** - users can simply drag and drop to access a link and drill down between files.

- **Prototype the Data Warehouse by creating** multi-dimensional databases on the client PC with dynamic regrouping at the click of the mouse.

- **Ability to share data** with other PC tools such as spreadsheets.

- **Windows 3.1** support.

LANSA/Server- unlock the power of your AS/400

LANSA/Server is communications middleware, providing lightning fast access to AS/400 data stored in the Data Warehouse.

- the **LANSA Repository** is used to transform normalised and unnormalized data into a user friendly Data Warehouse.

- **Enterprise Information Systems** can be easily built and customized using any front-end GUI tool that can call a DLL. Visual Basic programmers can concentrate on what they do best - defining the GUI using LANSA's friendly definitions. Your AS/400 data is protected by LANSA's security.

- **LANSA/Client** uses LANSA/Server to rapidly access AS/400 data.

LANSA Data Warehouse services

A range of services are available from your local LANSA agent to support the implementation of effective Data Warehouse for business user data access:

- **File Simplification Services**

  Use LANSA's Repository to simplify your existing system with its embedded arrays, internally described data and multi-member files, all without impacting existing applications. Business users can then easily access the simplified files.

- **Data Warehouse Design and Implementation**

  This service analyses your existing transactional system and business data access needs to propose an effective Data Warehouse design for business user data access, including identification of the most suitable pilot subject area, creation and maintenance of summary files triggered by operational updates, and proposed implementation plan.

- **LANSA/Client Administrator's Workshop**

  Administrators are responsible for creating easy-to-understand definitions of files and fields and automatic access routes between files. This course also covers the design of a Data Warehouse and the distribution of generated LANSA/Client applications.

This is only the Data Warehouse subset of the LANSA services available. Please contact your local LANSA agent for a complete list of LANSA products, courses and documentation available.

LANSA business user services

A range of services are available from your local LANSA agent to support the implementation of effective data access for business users:

- **LANSA/Client Business User Workshop**

  This course teaches the business user how to use LANSA/Client. It covers the basics of defining and executing queries, designing reports and creating charts.

- **LANSA/Client Advanced Business User Workshop**

  The advanced class teaches advanced reporting capabilities such as customized grouping and alerts.

- **LANSA/Client Private Tutor**

  An on-line tutorial referenced as needed when using LANSA/Client. It contains explanations and screen capture demonstrations of LANSA/Client features.

This is only the business user subset of the LANSA services available. Please contact your local LANSA agent for a complete list of LANSA products, courses and documentation available.