



D W Q
Foundations of **Data Warehouse Quality**

National Technical University of Athens (NTUA)
Informatik V & Lehr- und Forschungsgebiet Theoretische Informatik (RWTH)
Institute National de Recherche en Informatique et en Automatique (INRIA)
Deutsche Forschungszentrum für künstliche Intelligenz (DFKI)
University of Rome «La Sapienza» (Uniroma)
Istituto per la Ricerca Scientifica e Tecnologica (IRST)

M.A. Jeusfeld, T.X. Bui

Interoperable decision support system components on the Internet

Proc. 5th Workshop on Information Technologies and Systems (WITS'95),

Amsterdam, December 9-10, 1995

also appeared as Aachener Informatik-Berichte 96-17, 1996.

DWQ : ESPRIT Long Term Research Project, No 22469

Contact Person : Prof. Yannis Vassiliou, National Technical University of Athens,
15773 Zographou, GREECE Tel +30-1-772-2526 FAX: +30-1-772-2527, e-mail: yv@cs.ntua.gr

Decision Support Components on the¹ Internet

Manfred A. Jeusfeld (*), Tung X. Bui (+)

*) RWTH Aachen, Informatik V, Ahornstr. 55, 52056 Aachen, Germany
jeusfeld@informatik.rwth-aachen.de

+) Hong Kong University of Science & Technology, ISMT,
Clear Water Bay Rd., KLN, Hong Kong, tbui@usthk.ust.hk

Abstract: Operational software like online transaction processing systems is the lion's share of business information systems. Though many Decision Support Systems (DSS) were developed, they failed to become mainstream. The main reason is lack of knowledge of availability, applicability, and integration. In this paper, we propose to exploit the vast resource of the Internet as a second chance. Contrasting to other approaches, our strategy is to let users and developers of DSS's cooperate in weaving a web of distributed components. Repositories serve for ensuring application-centered metadata for searching. The method is scaleable in number of processes, domains, and platforms. A script language for combining DSS components is introduced based on observations from a real life example of a company.

¹ Most of the work reported here was done while the authors were with the Hong Kong University of Science & Technology, Department of Information and Systems Management, Clear Water Bay Rd., KLN., Hong Kong.

1. Introduction

According to a recent survey by PCWeek (1994), the deployment of Decision Support Systems (DSS) across organizations is rather limited. The reasons that explain this poor distribution of DSS:

- Difficulty in matching problems to appropriate DSS: Decision makers experience difficulties in finding available analytical models appropriate for their problems.
- Up-to-date DSS are not easily accessible: Comprehensive model-base is hard to maintain as models and tools are being continuously updated or improved.
- DSS are often too application-specific: Models have different assumptions about input/output formats; some of these assumptions are implicit, not explicitly declared format.

Recently, with the explosion of Internet use, many academic researchers started making their DSS components available to potentially interested users. Bhargava et al. [95;1] cite three factors contributing to the limited DSS deployment that, due to the currently limited size of the market for decision support systems, is a market potential. DSS deployment on the World Wide Web (WWW) is important, DSS availability on the WWW could eliminate, or at least reduce, the problems related to the issue of "re-investing in the hardware management. Goulal et al. [95] view DSS deployment on the Internet as an innovative means of providing opportunities for using newly developed DSS, (ii) validate DSS, in order to get feedback on their usefulness and performance, and (iii) promote widespread use of DSS via training.

From an operational point of view, the goal is to match the demand for DSS with much lower overhead costs than those of a more conventional distribution. Ideally, a decision maker should be able to combine DSS components in the same way that he/she builds and uses functions in a spreadsheet without having to be concerned with installation of the components in a specific environment. An Internet-based usable system must provide answers to the following questions:

- How can a complex decision problem be planned using distributed DSS components as an Internet-based approach to model integration? Our answer is to provide the user with a graphical and a textual representation.
- How can users access appropriate DSS components dispersed in an unlimited Internet search space? Here we use the idea of Uniform Resource Locator (URL) known from the World Wide Web. Data, scripts, models, and DSS components are all identifiable by their URL.
- How can data objects be transformed into the format expected by DSS components? We propose a version of the uniform data exchange OEM which provides a standard information with the complex DSS data. This minimizes the hand-off between sender and receiver of DSS data and allows to store a data item in its original format.

- How can the Internet traffic be minimized when data/script objects are distributed across nodes? Here, the simplicity of our script language will allow for optimizations.

Our approach is characterized by a lack of central control. DSS components (models, scripts, programs) are created and maintained in a distributed manner. The sum of all available DSS components may not have a consistent model. Therefore, we propose repository systems mediating between DSS processes. These repository systems define which components may be combined and how to browse in an application-specific search space of components.

A Decision Support Assistant (DSA) was proposed by [Felter 89] for local networks (LAN). It provides an intelligent assistant that helps a user select DSS components from a model base. Felter studied interference of concurrent DSS invocations and noticed the limited processing time resources. In our approach, one of the DSS components is more complex than that in a LAN. The main difference is the degree of distribution. In Felter's approach, all DSS components are managed by a central DSA. In ours, a federated environment makes a central controller and undesirable. In that regard, we depart from the 'brokers-agent' model [Bhargava et al. 95;2]. Potential DSS users do not need to buy the software; they make use of distributed computational resources on the Internet.

To allow distributed management of all resources (DSS script and data), we present two design principles: A uniform data representation as well as exchanging information between heterogeneous systems and identifying input/output data as well as the DSS components that participate in a process. In addition, we advocate the implementation of repository systems for DSS users with the knowledge to design distributed decision support systems. The installation of DSS components across the Internet. As a whole, our approach fits into the framework of the "enterprise information bus" by [Bauer et al. 95] with the requirement of having a central agent.

To illustrate the approach set in this study, a case study adapted from real-life application is presented. International (RI) is a multinational company with its headquarters in North America. With annual sales of Can\$ 1.5 billion, RI has its operations in a wide variety of products offered, RI seeks to constantly monitor its market performance with its geographically distributed sales, RI information system is distributed with major computing hubs located in Montreal, Canada and Paris, France, as well as various regional computing centers, e.g., in Frankfurt, Germany.

As part of its decision support and executive information system, RI has developed a financial planning system. The purpose of the system is to establish

- yearly and 5-year budgeting
- short term and long term strategic planning
- cash flow management
- routine and ad hoc data analysis and reporting

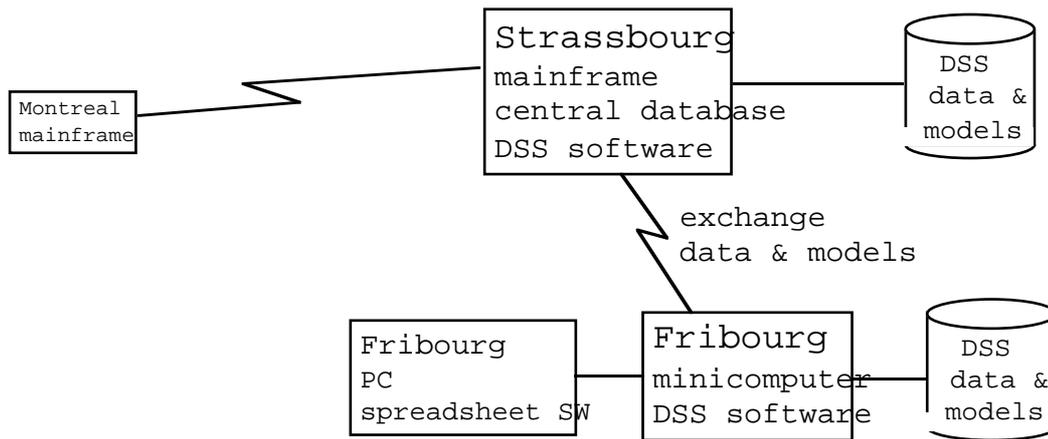


Fig. 1 Architecture of an existing federated DSS

Figure 1 shows an excerpt from the current federated DSS architecture. European headquarters in Strassbourg maintains a financial data company's European operation. Regional centers, like Fribourg in Switzerland, provide data to the headquarters but also require information for decision making. Some financial data and models are kept locally at both the headquarters and the regional computer centers. Both provide DSS software for evaluating the financial models on the financial data. Spreadsheets on personal computers serve for visualizing results of the analysis. This federated DSS architecture has been implemented by leased data communication protocols. The location of the financial data is determined in group meetings between the headquarters staff and regional staff.

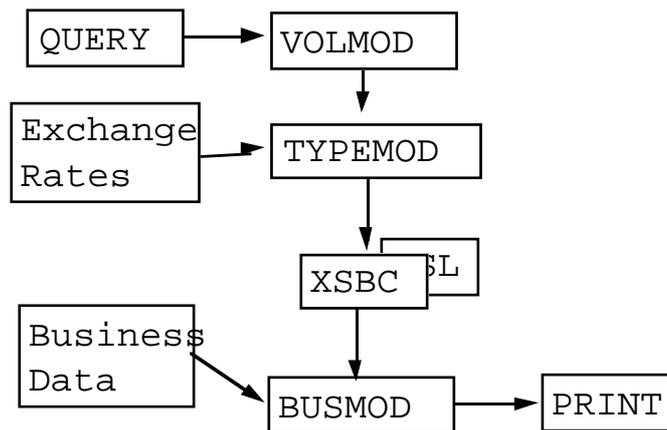


Fig. 2 A script for financial analysis (from RI)

Scripts for computing an analysis were specified by semi-formal diagrams (see Fig. 2) and then implemented by a 4GL language. The boxes on the left represent financial data, the boxes in the middle are financial models, and the boxes on the right represent how the output is treated. RI also has examples of scripts with local analysis for all months of a fiscal year.

The example shows that knowledge about the DSS is spread in heterogeneous (data, models, scripts). Interoperability, i.e. the ability to process without information loss, is ensured by using the same software dedicated transformation routines (wrappers). Scripts have a Janus-like nature: they can be viewed as a high level programming language, on the one hand, and as a semantic relationship between data and models, on the other.

2. Overview of the Proposed Federated System

We use the setting of RI as a running example for our method. The components of the system are distributed on the Internet. Our goal is to provide a scalable method where DSS components and processors may be located anywhere on the network. Users (decision makers, DSS software providers, etc.) of the proposed federated architecture can exchange information after an agreement on location, representation, and access. The information has been reached. The users may be in different departments of the same company but may also be from different cooperating companies.

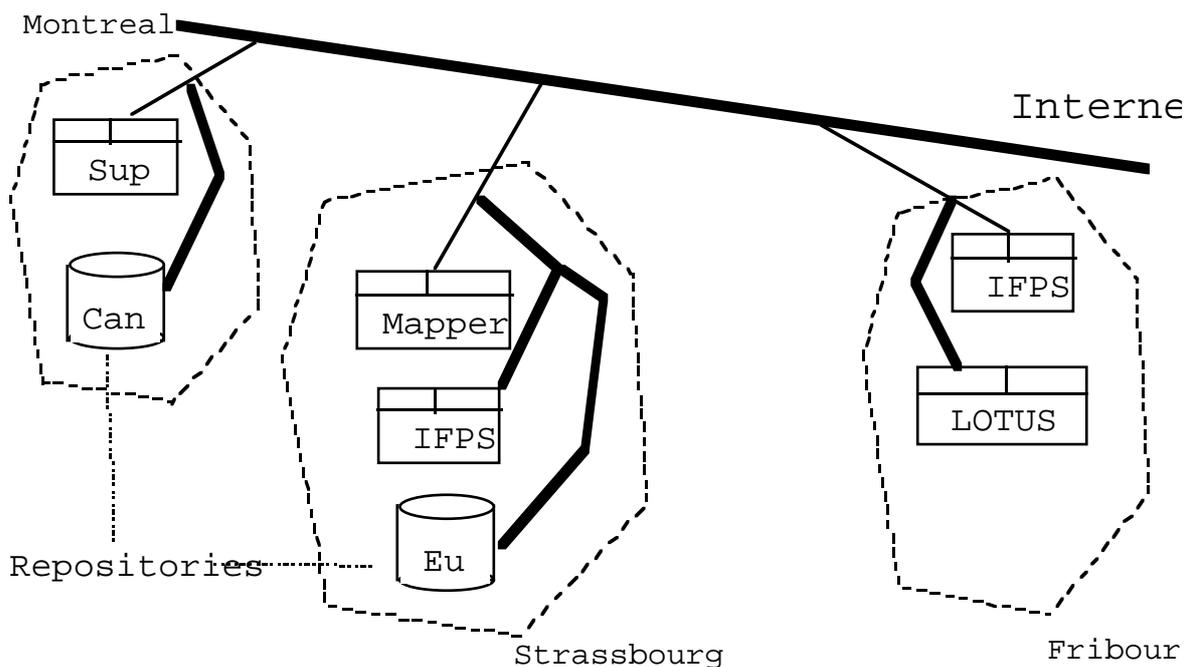


Fig. 3 Overall architecture of the proposed federated DSS

Figure 3 shows a simplified view of a network of DSS processing systems (labeled 'Can' and 'Eu' in the Figure) connected by the Internet (thick lines). The rectangular nodes represent DSS processors, i.e. systems that provide information relevant for some decision. The cylindrical nodes (here 'Can' and 'Eu') are examples of repository systems. The repositories constitute the foundation of the federated system in that they serve as the main information resource to support the composition of Internet-based DSS components. The repositories can be

standard database application program, or by knowledge-based systems in a specific application domain or purpose (e.g., repository of functions for statistical procedures, etc.).

Each repository creates its own name space, i.e., the set of names for script and data objects. Specialized name spaces are justified because different groups of decision makers are interested in different DSSs and more importantly, distributed repositories create a higher degree of security: script and data objects are known only to those users who have access to them.

One of our goals is to help DSS providers promote the access to their systems. The overall system can be continually updated by installing new components.

1. A user (system administrator or even decision maker) browses a repository for DSS processors suitable to her problem definition, and selects one or more.
2. The repository is queried for the installation procedure of the selected processor(s).
3. The installation procedure is downloaded to the local system and checked to see if the features of the procedure conforms to the local system configuration.
4. At the end of a successful installation, script objects for calling the DSS processor have to be inserted in the local repository (uniformly named, etc.).

As more and more copies of the same DSS package are spread over a network, redundancy can be exploited for optimizing the load balance of the system. A DSS provider 'publishes' a new DSS component by inserting its name in the repository system known to him. This can include a test installation of the component by the DSS provider to help potential users test the component(s).

3. Interoperability Techniques

Let's consider the following generic scenario. A DSS component P1 produces a data object O1 which is later used by DSS component P2 as input. They are autonomously developed and the data exchange via O1 may not have been foreseen at the time P1 and P2 were designed. The traditional approach is to store O1 in a database for later processing (e.g., [Sprague 1989]). For this purpose, the database must be known in advance and is relatively fixed. In our approach, data objects as script objects are assumed to be distributed and no common type system is required.

Remember two principles for distributed processing: first, data must be converted into a transferable exchange format similar to IDL [Nestor et al. 1994]. Resource Locators (URL) [Barners-Lee 95] can be used to identify the location of objects installed on the Internet.

Uniform Data Representation. We regard any data as a string of characters received and interpreted by the DSS component which processes it. No external knowledge about the kind of data, and is therefore not suitable for environments where data and script objects evolve without central control.

There are several similar approaches from the areas of distributed heterogeneous databases, and distributed artificial intelligence cooperation data exchange. We adopt the OEM format [Papakonstantinou et al. 95] to represent the type information within each data object, i.e., the data structure of the object representation. Objects are represented as terms (n, t, v) where 'n' is the object, 't' is its type and 'v' is its value. The value is either a scalar or a term. As an example, consider the definition of a vector of currency values. The building principle of OEM objects are triplets (instance-name, type-name, value) where the first component identifies the object, the second is the name of the type and the last is the value. OEM objects may be arbitrarily nested. Nested objects are identified by a path expression, e.g., 'Selling-Data-1995.Q-2' identifies the second component of the 948 (in thousand). Such a representation makes a data object independent of the programming language: it simply carries its type definition. A system that reads such a term can extract the type information and match it with its own system.

```
(Selling-Data-1995, Selling-Vector,
  ( (Q-1, K-CAN$, 1033),
    (Q-2, K-CAN$, 948),
    (Q-3, K-CAN$, 1018),
    (Q-4, K-CAN$, 316) ))
```

Fig. 4 OEM object of type matrix

The input of DSS scripts are data objects that may be located at one or more (remote) locations. Therefore, both data objects and their locations are identifiable across system boundaries. Uniform resource locator (URL) [RFC Consortium 95] serve this purpose. The general format is

```
protocol://node-address/path
```

The component 'protocol' specifies the communication protocol to be used to retrieve data or to activate a remote process. We assume the widely-used protocols are listed in the following. The 'node-address' uniquely identifies the computer system where the requested data or script object is located. Finally, the 'path' identifies the location in the space of the local computer system. For example, the URL

```
http://strassbourg.rubber.fr/finance/
  Selling-Data-1995.oem
```

uniquely identifies location of the data object in the file system of the computer 'strassbourg' of organization 'rubber' in France. Via the URL, the data is technically accessible from any node in the network in which 'strassbourg' is located. We assume that any persistent data object, i.e., a data object that is not processed during the processing time of the script objects that produced or consumed it, is stored in a file with extension 'oem' indicating that its content is in OEM format.

A call to a DSS processor is also named via URLs. The format of a procedure call with input parameters referenced by their URLs. For example,

```
http://fribourg.rubber.ch/bin/IFPS?call=http://
strassbourg.rubber.fr/scripts/Earning-Forecast-1995.oem
```

invokes the DSS processor IFPS (actually a financial analysis package) to evaluate on the input object Earning-Forecast-1995.oem located in the directory specified in the HTTP protocol, such an URL names the result of the call to be exploited in the definition of script objects below.

Figure 5 shows a script object where all objects reside on the Internet. The script object Earning-Forecast-1995.oem encodes the call to the processor IFPS at Fribourg, and the arguments of the call which are other script objects or calls of component scripts. Argument arg-1 of Earning-Forecast-1995.oem is the financial model BUSMOD, arg-2 is Selling-Data-1995 from Strassbourg, and arg-3 is the result of a nested call to another DSS processor IFPS at Strassbourg.

```
(Earning-Forecast-1995, po-call,
  ((po-node,url,http://fribourg.rubber.ch/bin/IFPS),
   (arg-1,url,
    http://fribourg.rubber.ch/local/BUSMOD.ifps),
   (arg-2,url,
    http://strassbourg.rubber.fr/finance/Selling-Data-
    1995.oem),
   (arg-3,po-call,
    ((po-node,url,http://strassbourg.rubber.fr/bin/IFPS),
     (arg-1,url,http://strassbourg.rubber.fr/local/XSBC.ifps),
     (arg-2,url,
      http://strassbourg.rubber.fr/local/Exchange-Rate-
      1995.oem))
    )))
```

Fig. 5 OEM representation of a script object

The evaluation of a script object is explained in section 5. It is sufficient to note that the script object encodes a data flow diagram. Also note that a script object is in OEM format like the data objects. It can be the output of a compiler script. Furthermore, it can be class presented in section 4.

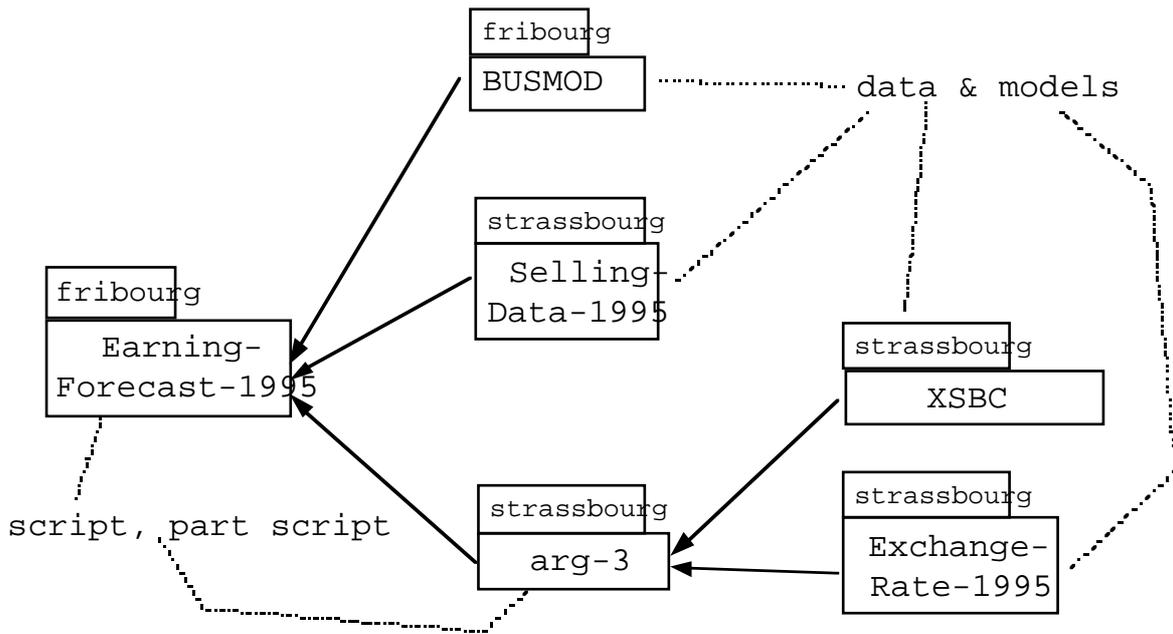


Fig. Diagram denoting a distributed DSS script

The URL conventions are sufficient to invoke a remote DSS component or remote input objects and to receive the result in as a single Legacy DSS components can be integrated by a transducer [Geneseret which mediates between the network and the DSS component. The trans call in the above syntax, transforms the (input) data into the inter legacy DSS component and then invokes it.

The proposed approach is well-suited for DSS components interaction, i.e., those that are linked with other DSS components. interface should not impose principal obstacles (i.e., at the client framework). An elegant way is to install a specialized DSS processor system of the user. This processor accepts requests from the remote the URL call mechanism shown above. Locally, it gathers the input appropriate input device.

4. The Repository

Uniform data representation and naming described above provide a framework for communication between distributed DSS components. repository which stores conceptual information about data objects and answers queries based on classification hierarchies.

Persistent Data Objects in the Repository object is in OEM format. The type can be extracted from its representation and be used for repository. We use the statement 'OBJECT x IN T' to express that instance (element) of type 'T'. The type is represented as a term or and base types (given by their name). For example, the most specific

1995 is Selling-Vector(K-CAN\$,K-CAN\$,K-CAN\$,K-CAN\$). Consequently, object is represented in the repository as:

```
OBJECT http://strassbourg.rubber.fr/finance/  
        Selling-Data-1990.oem  
IN Selling-Vector(K-CAN$,K-CAN$,K-CAN$,K-CAN$)
```

Membership of the data object to more generic types like Selling-Vector(Int,Int,Int,Int) can be derived provided Int is a more generic type. We express such generalization in the repository by a statement like Selling-Vector(K-CAN\$,K-CAN\$,K-CAN\$,K-CAN\$). Further classifications can be obtained when additional attributes are provided. Such attributes are highly domain-specific, especially when they come from different organizations. Therefore, it is difficult to prescribe a standard set of attributes to be used. Book-keeping features like 'creator', 'creation-date', 'valid-thru' are reasonable first choices. Semantic information explaining the meaning of the problem class and applicable models is more useful for data classification than common terminology among the users of the repository (which is beyond the scope of this paper). In our running example, the attributes could be as follows:

```
K-CAN$ ISA Int  
...  
OBJECT http://strassbourg.rubber.fr/finance/  
        Selling-Data-1995.oem  
IN Selling-Vector(K-CAN$,K-CAN$,K-CAN$,K-CAN$)  
WITH ATTRIBUTE  
    creator: "Francois Nome";  
    creation-date: "4-Nov-1995";  
    valid-thru: "31-Dec-1995";  
    remark-1: "entries correspond to quarterly selling  
              of Rubber-Europe";  
END
```

DSS-Scripts are special data objects that describe a computation problem. In the example below, the script Earning-Forecast-1995 is defined with arguments of type IFPS-Model (a plain text) or Selling-Vector, respectively expected to be of type matrix(Int). Note that the input and output of the script are searched for data objects matching a script's signature and conformance.

```

OBJECT http://strassbourg.rubber.fr/scripts/Earning-
Forecast-1995.oem
IN DSS-Script
WITH ATTRIBUTE
    arg-1: IFPS-Model;
    arg-2: Selling-Vector({Int});
    arg-3: Exchange-Vector({pair(Real,Real)})
    result: matrix(Int)
    preferred-processor:
http://fribourg.rubber.ch/bin/IFPS
END

```

The processors of scripts (attribute preferred-processor in the ex software packages installed on some computer in the network. It m same DSS package is installed on several nodes. Then, there is a ch strategy of the script. In the repository, we store a statement like

```

OBJECT http://fribourg.rubber.ch/bin/IFPS
EQUIVALENT-TO
http://strassbourg.rubber.fr/software/bin/IFPS

```

to represent this fact. It is used for the run-time scheduling of th

Some scripts may have the only task to filter, transform or objects. They constitute no special case. For example, the result o may be transformed into a spreadsheet format and then displaye spreadsheet package.

5. Implementation Steps

To implement the proposed approach, two principle steps are r DSS components have to be connected to the Internet, and second , t in the repository for identification.

Connecting DSS Components to the Network. Our approach relies protocols and interfaces. Therefore, setting up a running system d amount of programming. First, the DSS processors must be accessible This can be achieved by installing a so-called HTTPD server [WWW-C the local node. The server has the basic function of accepting cal returning some result (in our case data objects). It can be config with the right URL to the local DSS component. Second, a local DSS able to retrieve the arguments of a call (as encoded in the script remote HTTPD servers. The standard-compliant library 'libwww' for available from [WWW-Consortium 95]. Figure 7 shows a schematic v component linked to the WWW.

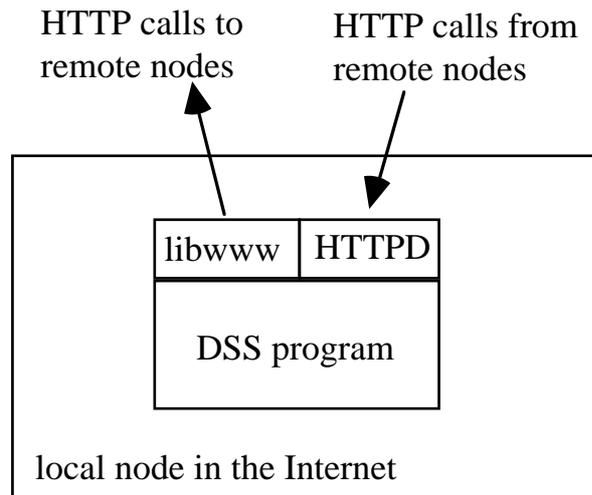


Fig. Schematic architecture of a DSS processor linked to WWW

This architecture allows a user to activate a DSS processor from a certain script object. As third and final requirement for the connection the result of a call to a script are expected to be in OEM format the processor must be able to read/write such a format and transfer the internal format. The effort for this function is minimal once the OEM script is invoked by submitting a call like

```
http://fribourg.rubber.ch
    /bin/IFPS?call=http://strassbourg.rubber.fr
    /scripts/Earning-Forecast-1995.oem
```

to the HTTPD server of a suitable processor. The DSS processor the object (here Earning-Forecast-1995.oem) through its libwww interface descriptions in the script object which are of type 'po-call', for example stored on a local file. e.g. http://fribourg.rubber.ch/temp/arg-3.oem. The DSS processor can issue the corresponding call, in our running example

```
http://strassbourg.rubber.fr/software/bin/IFPS?
    call=http://fribourg.rubber.ch/temp/arg-3.oem
```

in the same manner it was called itself. The local DSS processor can deviate from the preferred processor specified in the script and call a local processor, e.g. the IFPS package at the local site. This scheduler control and is guaranteed not to influence the result of the computation. DSS processors must be equivalent. Finding equivalent processors is a (local or remote) repository.

Searching the Repository. The main functions of the repository system are to manage the properties of data and script objects as shown above, and to facilitate the search for information about data and script objects. Three search strategies are

Simple search for ~~scripts~~ ^{Assume}. A decision maker has a problem encoded by an input data object and wants to find a suitable collection of DSS components. The search can be determined by matching the type of the data object with the types of the DSS scripts stored in the repository. The search may be limited by the features of the DSS script. The search is efficient, more exactly, the number of objects in the repository. Indexes on the input argument features provide an average case performance to constant time.

Chain search for ~~scripts~~ ^{Assume}. Assume that the repository contains some data transformation scripts P_1, \dots, P_k such that the input argument of P_1 matches the input argument of P_2 , the output of P_2 matches the input of P_3 , and so on, until the output of P_k matches the input of P_{k+1} . Then, suitable solutions can be obtained by searching such chains and then finding a DSS script whose output matches the output of P_k . This search delivers more alternative solutions than a simple search. Depending on the number of data transformation scripts known to the decision maker, the search is linear in the number of matches between input and output arguments of components in the repository. In the 'worst' case (i.e., every input argument is transformed to every output type), it is quadratic in the number of matches. The quality of the search, and subsequently, the problem solution depend on the set of transformation scripts.

Interactive script object generation. This is both the most useful and the most complex search. It assumes that the decision maker has more than one data transformation script definition and that some goal objects have to be computed by a combination of scripts. Theoretically, all combinations of chains, each starting with an input data object, are applicable. Indeed, the number of combinations explodes as the number of input data objects and available scripts increases. The problem of finding a suitable chain of objects is similar to scheduling problems. Since the repository has a large number of available script objects (only input/output types plus some other properties), the goal of the decision maker, we propose an interactive approach to generate a script object. The user incrementally builds the script starting from an input data object. Simple and chain search are used as auxiliary functions which offer a choice of possible component scripts operating on a selected data object or output of a script object). After validation the generated script object is stored in a repository for later use and re-use (then being component of other chains).

Implementation Properties. The script objects encode functional expressions in the OEM format. The functions are denoted by the URLs of the script objects and the arguments are either functional expressions or URLs of data objects. The following computational model for DSS processors:

Persistent data objects are never updated after their creation. That means that the OEM term referred to by an URL is always the same as when it was created. This is analogous to the treatment of variables in functional and logic programming.

is different from the destructive updates allowed in imperative programming. The result of a script object only depends on its input objects. As of the same DSS processor will have the same result if and only if persistent data objects.

The computational model allows application of several heuristics of script objects.

- Lazy evaluation lets arguments be evaluated only when actually needed. This heuristic reduces the overall load on the network. Its implementation is difficult because it affects the internal routines of the DSS components.
- Caching materializes results of sub-expressions (parts of script objects of type 'po-call' for speeding up subsequent calls of the same sub-expression). This makes most sense for script objects having only persistent data objects whose result never changes.
- Multi-processing allows arguments of a script object which are of type 'po-call' to be called in parallel. This strategy is the opposite to lazy evaluation. The response time provided the calls go to distant idle nodes in the network. The main problem for this strategy is that the parallel calls do not interfere with each other. This is achieved by forbidding updates to persistent data objects.
- Load balancing can be enforced to optimize network traffic. Calls to script objects (named by their URL in the script object) can be replaced by calls to script objects located elsewhere. Load balancing can be incorporated into the DSS compiler (compile-time approach) or during evaluation (run-time approach). Multiple processors may be installed redundantly at any time on any node in the network. The overall system is scaleable to the workload. There is no need of heuristics to optimize network traffic. The (limited) knowledge of the local network topology can be applied with the (limited) knowledge of the local network topology.

Script objects are hierarchical, i.e., they do not contain local objects. The evaluation of a script object will terminate if the evaluation of its components (elementary script objects) terminate. We conclude the following property:

The evaluation of script objects will always terminate under the lazy evaluation strategy (e.g., top-down, depth-first).

This property allows DSS users not familiar with programming languages to use DSS without fearing to consume too much resources of the network. In fact, the amount of resources consumed can be estimated from a script object definition. The total processing time is the sum of the processing times of the component scripts in the script object. This is linear in the size of the script object. The processing times of the component scripts are in general indefinite. Response time, i.e., the time between the request for a script object and receiving its result, depends on the degree of multi-processing. The response time is the processing time divided by the number of distinct processors. The actual response time also depends on data dependencies and network bandwidth. Network costs (measured in the amount of data transferred between distinct nodes) is also linear in the size of the script object.

only as many transfers occur as there are DSS processors referred to. Network costs are also linear in the size of the input data objects. Any call to a DSS processor is linear in the size of its inputs. This is a reasonable assumption for DSS components since their purpose is to reduce large data sets to a small number of choices. Network costs can be reduced by installing any DSS processor referred to in the script object on the local Internet. This is an extreme case of load balancing whose expenses are borne by local resources.

Integrity and Control. At first sight, the distributed approach on the Internet is in danger of uncontrolled access to models which were never meant to be shared. The reader should however note, that the repositories are specifically designed to be a space of meaningful data and scripts objects for a certain group of users. In any case, a repository can just cover the objects on a local computer. Even small organizations have worldwide connections and will be able to provide for global decision support. The repository for such an organization is not restricted to cover those distributed objects important for its business.

As seen for the application of electronic data interchange (EDI), the application of information processing requires a clear understanding of the business and the participating organizations. We assume that this understanding is the basis for the exchange of models and software. The 'global' DSS, i.e. the sum of all DSS objects on the Internet, can never be assumed to be consistent. Instead, it consists of which consistent subsets (maintained in the decentralized repositories). This global resource will increase in value when more and more distributed DSS objects and scripts are published in the repositories. They are in fact designed for interoperability.

6. Final Remarks

A combination of three elements describe our approach. Uniform data objects and data objects allows the programs and data to be located anywhere. A uniform data format, OEM, allows data exchange independent from the language of the DSS components. Finally, decentralized repositories, distributed scripts and the evolution of the overall system by semantic information.

The proposed approach relaxes the necessity of having a central clearinghouse for the overall system. While the lack of a central clearinghouse creates some problems in problem formulation, book-keeping issues, our approach empowers a large number of decision makers (users) and DSS providers to participate in the system. It also provides more flexibility in terms of DSS proliferation. Utilizing widely accepted protocols for data exchange keeps the network costs reasonably low. Also, the effects of low bandwidths of existing Internet are reduced by duplicating DSS processors on the nodes (computers) where needed and a large number of nodes can be exploited via multi-processing.

Future efforts should concentrate on more sophisticated search repositories and quality control of retrieved DSS scripts. Two other: cooperation among several decision makers and guidance in long-term The approach presented in this paper has only marginal support f makers cooperating in a decision script. To extend our approach to has to embed the individual script objects into complex script or p & Jarke 92]. Then, loops for scripts are likely to become an iss: investigating such limited forms of such loop concepts.

References

- [Ba et al. 95] S. Ba, Kalakota, A.B. Whinston, "Executable Documents DSS", Proc. 3rd Intl. Conf. on DSS Hong Kong, 22-23 June 1995, pp.
- [Barners-Lee 95] T. Barners-Lee, Uniform Resource Locators, Document http://www.w3.org/hypertext/WWW/Addressing/URL/URL_TOC.html, 1995
- [Bhargava et al. 95;1] H. Bhargava, A. King, D. McQuay, "DecisionNet for Modeling and Decision Support over the World Wide Web", Proc. 3rd Intl. Conf on DSS Hong Kong, 22-23 June 1995, pp.499-505.
- [Bhargava et al. 95;2] H. Bhargava, R. Krishnan, D. Kaplan, "On Gene WWW-based Network of Decision Support Services", Proc. 3rd Intl. Hong Kong, 22-23 June 1995, pp. 507-515.
- [Ellis & Jarke 92] C. A. Ellis, M. Jarke (eds.), Distributed Cooperative Information Systems, Proc. 3rd Intl. Workshop on Intelligent & Cooperative Information Systems, Report AIB-92-18, RWTH Aachen, Germany, 1992
- [Felter 89] R. Felter, Decision Support Assistansys, Deutscher Univer Wiesbaden, Germany, 1989.
- [Genesereth & Ketchpel 94] M.R. Genesereth, S. P. Ketchpel, "Software Communications of the ACM, 37, 7, 1994, pp. 48-53 & 147.
- [Goul et al. 95] M. Goul, A. Philippakis, M. Kiang, D. Fernandes, B. Client/Server Open-DSS Protocol Suite for Automating DSS Deployment World Wide Web", Proc. 3rd Intl. Conf. on DSS Hong Kong, 22-23 June 1995, pp. 517-528.
- [Nestor et al. 90] J.R. Nestor, J.M. Newcomer, P. Gianni, D.L. Stone and its Implementation, Prentice Hall, 1990.
- [Papakonstantinou et al. 95] Y. Papakonstantinou, H. Garcia-Molina, "Information exchange across heterogeneous information sources", Proc. 11th Int'l Conf. on Engineering, Taipei, Taiwan, 6-10 March 1995, pp. 251-260.
- [Rasmussen et al. 91] J. Rasmussen, B. Brehmer, J. Leplat, Distributed Cognitive Models for Cooperative Work, John Wiley & Sons, 1991.
- [WWW-Consortium 95] WWW-Consortium, CERN Httpd, Document <http://www.w3.org/hypertext/WWW/Daemon/Status.html>, 1995.