Article

Design Limitations, Errors and Hazards in Creating Decision Support Platforms with Large- and Very Large-Scale Data and Program Cores

Elias Koukoutsis 1,*, Constantin Papaodysseus 1, George Tsavdaridis 2, Nikolaos V. Karadimas 3, Athanasios Ballis 4, Eirini Mamatsi 1 and Athanasios Rafail Mamatsis 1

1 School of Electrical and Computer Engineering, National Technical University of Athens, 15708 Athens, Greece; CPapaod@cs.ntua.gr (C.P.); eirinimamatsi@gmail.com (E.M.); trmamatsis@mail.ntua.gr (A.R.M.)
2 Hellenic Army General Staff, 15708 Athens, Greece; GTsav@central.ntua.gr
3 Division of Mathematics and Engineering Science, Department of Military Science, Hellenic Army Academy, 21100 Nafplion, Greece; nkaradimas@sse.gr
4 Department of Transportation Planning and Engineering, National Technical University of Athens, 15708 Athens, Greece; abal@central.ntua.gr

* Correspondence: E.Koukoutsis@ece.ntua.gr

Received: 26 October 2020; Accepted: 10 December 2020; Published: 14 December 2020

Abstract: Recently, very large-scale decision support systems (DSSs) have been developed, which tackle very complex problems, associated with very extensive and polymorphic information, which probably is geographically highly dispersed. The management, updating, modification and upgrading of the data and program core of such an information system is, as a rule, a very difficult task, which encompasses many hazards and risks. The purpose of the present work was (a) to list the more significant of these hazards and risks and (b) to introduce a new general methodology for designing decision support (DS) systems that are robust and circumvent these risks. The core of this new approach was the introduction of a meta-database, called teleological, on the base of which management, updating, modification, reduction, growth and upgrading of the system may be safely and efficiently achieved. The very same teleological meta-database can be used for the construction of a sound decision support system, incorporating elements of a previous one at a future stage.

Keywords: very large-scale decision support systems; very large-scale data and program cores of information systems; meta-database; teleological meta-database; thematic list; indicators list; computational methods list; geographically dispersed systems; external sources

1. Introduction

Nowadays, there is a considerable number of public and private operators, who make extensive use of decision support systems (DSSs or DS systems) in a surprising large number of operational procedures and businesses. These operations include:

- Vehicle or vessel fleet management [1–3];
- Marine services and policies (testing, inspection and certification services in the areas of quality, health and safety, as well as security, environmental considerations and relevant policy impacts, etc.) [4–6];
- Management of goods in road, rail, sea and intermodal transportation. These operations are very complex, involving public and/or private train, truck and ship owner companies, goods
owners and buyers, transport operators, intermodal terminal operators, infrastructure providers, safety and security operators, etc. [7–9];

• Handling of emergencies and crisis [10–13];

• Complex engineering tasks and, in general, large and complex sequences of operations and/or tasks covering very extended geographical areas (e.g., aircraft and sea vessel design and very large and/or complicated production lines in factories and many more) [14–17];

• Transport planning (estimation of important parameters than cannot be measured, forecasts, policy making, etc.) [18–21];

• Cost-benefit analyses [22–24];

• Healthcare, a domain that is fundamental for human society, especially nowadays (in 2020) due to the COVID-19 pandemic [25–27].

These examples of operations constitute only a small number of those encountered in practice. In addition, in most cases, each associated DSS employs sets of very complex computational models and algorithms. In essence, a DSS is a special kind of a very complex information system (IS), where an information system is an integrated set of components for collecting, storing and processing data and for providing information, knowledge and digital products [28].

The present manuscript is organized as follows:

In Section 2, the authors deal with the recent tendency/necessity to build very large-scale decision support systems. In fact, it is argued that the advances in information and communication technology have given the opportunity to engineers to try to develop DSSs that tackle very big and/or particularly complicated practical problems.

In Section 3, a number of serious difficulties and problems are stated, in connection to specific properties/characteristics of a very large-scale decision support system. These properties concern the following:

• The ability to efficiently update a DS system;

• The flexible expansion of the DSS;

• The consistent reducibility of an information system;

• The ability to make changes to the information system;

• The upgrading of the DSS.

In Section 4, crucial characteristics of contemporary, mainstream methodologies for creating large-scale decision support systems are reported. It is argued that any ad hoc development of a DS system is bound to suffer from serious intrinsic problems, which only a consistent fully methodological approach can resolve.

In Section 5, the authors introduce a methodology that circumvents all the aforementioned problems. In Section 6, the authors state the reasons for which the introduced methodology resolves these problems associated with very large-scale DSS development.

Finally, in Section 7, the conclusions of the present work are stated.

2. The Tendency to Increase the Volume and Complexity of Data and Program Collection of the DSS

A considerable number of contemporary DSS applications eventually ask for very large data and program collections. The tendency to increase the size and complexity of these collections is due to a variety of factors, such as:

(a) Advances in information technology and communications (ITC) have created a market for DSSs with very large-scale data and program collections.

In fact, progress in software design and implementation makes possible the almost instant handling of huge data sets. For example, the conceptual development of balanced, N-ary trees allows
for accessing a specific entry among trillions of entries in four or five very fast steps only. In addition, contemporary hardware allows for creating and managing/handling “storages”, that is, devices that may contain thousands of disks (hard disks or DSSs).

(b) The same ITC advances have opened the way to engineers to attempt to develop DSSs that tackle very big and/or complicated practical problems. Two related examples follow:

- The effort for developing the European Transport Information System for Policy Support (ETIS).

Maintaining an efficient transport infrastructure and formulating a common transport policy are critical elements for the economic and social development of the European Union (EU). In 1994, the EU Commission adopted a comprehensive proposal for the “Trans-European Networks” (TEN-T) guidelines [29], which are plans for improving the performance and further development of the complex EU infrastructure. The first version of these guidelines covered 70,000 km of rail tracks (including 22,000 km of new and upgraded tracks for high-speed trains), 58,000 km of roads (including 15,000 km of new roads), corridors and terminals for combined transport, 267 airports of common interest, networks of inland waterways and many seaports. It is evident that these numbers have greatly increased since the adoption of TEN-T guidelines.

The effort for performance improvement of the Trans-European Networks and the application of a successful transport policy requires significant information concerning the current status of the networks and the corresponding inhabited regions. More specifically, this required information includes passenger and freight transport volumes, traffic congestion, environmental impacts of transport, etc. It is also necessary to consider reliable forecasts for the time evolution of the aforementioned factors. The relevant information exists in a large number (more than 500) of dispersed, heterogeneous and autonomous databases throughout the entire Europe. In addition, more relevant information is acquired from the results of an ensemble of pertinent scientific studies. The European Transport Information System for Policy Support has been envisaged as a tool for the support of decision making on transport policies and policies related to the Trans-European Transport Networks. ETIS must accommodate policy-related information in a repository that has to be kept up to date by experts. In ETIS, policy questions are linked to the relevant data sets of the sources via a hierarchical structure of policy criteria, policy indicators and data variables [20].

- A second example is the information system for the port of Rotterdam. This system includes a huge amount of geographical information. It is noteworthy that the port of Rotterdam incorporates more than 5000 docs, an airport, a considerably large storage area for alumina and abundant apron spaces for temporary storage of goods and containers.

The IS of the port of Rotterdam, called Port Community System (PCS), provides services that focus on all port sectors, such as containers, break bulk, dry bulk and liquid bulk. Anyone in the logistics chain can easily and efficiently exchange information through these services. Moreover, the PCS offers a package of properly tailored services to each of the target groups of clients and operators.

From the entire previous discussion, it is evident that the information involved in the PCS is huge and very complex [30].

(c) The inclusion and use of data, which are directly or indirectly georeferenced (map-related data), in a DSS. In the first case, the direct one, the exact co-ordinates of an object are explicitly provided. On the contrary, in the second case, the indirect one, there is a link to an object, which determines the data co-ordinates. For example, if an accident takes place at a specific point of a certain road and then, if the exact co-ordinates of the accident location are registered in an IS, this event is directly georeferenced. On the other hand, if the registered information is something like the following: “the accident happened at the 23rd kilometer of the road from city A to city B”, then the event is indirectly georeferenced.
Usually, the main bulk of the information is indirectly georeferenced. A surprisingly large part of the information that should be included in such a DSS is georeferenced, even though this georeferencing is well hidden at a first glance. For example, in most nations, such as the USA, the laws and directives may drastically change when crossing state boarders. Even in Greece, where the legislation is much more uniform, when one travels from an island to the mainland, one faces various changes in authority jurisdiction and legislation in the interior of the island, the mainland, at ports and on the sea (near-port area, national waters and international waters).

It is evident that the requirement that a large-scale decision support system handles and uses map-related information may greatly increase the volume of its data and program core.

(d) Finally, special DSS architecture requirements, like the following:

- Distributed and/or geographically dispersed systems, where properly selected duplication of part of the information and programs must be implemented, in order to increase the efficiency of the local systems. For example, a very frequently accessed part of information by all sub-systems must be preferably kept in every local sub-system.
- Umbrella-type systems that receive and process information in real time from a large number of heterogeneous, autonomous and, possibly, legacy systems. Sometimes, these systems may cover a large part or the whole of a continent. Such is the case of a European system for maritime surveillance, the complexity and size of which is evident.

3. Severe Difficulties Appearing in the Actual Deployment of a Very Large-Scale Decision Support System

We use the general term “large data collection” in order to describe the extended data core of a large-scale decision support system. The data collection may include dozens or even hundreds of groups of loosely related data sets or databases, the content of which may be thematically divergent.

In this work, we will often use examples of large-scale decision support systems for policy support and decision making in the area of transport. We will do this for two main reasons:

- In many transport DS systems, practically all problems referred to in the present work have already emerged, due to the huge size and complexity of the involved data. As a first example, we state the large number (more than 50) of Federated States in the USA. An analogous second example is the European Transport Information System, which is a DSS for supporting transport policy decisions in all EU countries.
- Engineers and other scientists that have developed the pioneering versions of the transport DS systems have faced a considerable number of serious difficulties and problems, many of which still remain unsolved. It must be emphasized that for a considerable number of these difficulties and problems, the severity of the resulting hazards may be anticipated.

We would like to mention the following, widely used terminologies in transport DS systems: The data and program collections are often called “repositories”, “observatories” or “transport observatories”. Frequently, when the transport observatories cover exclusively the needs of a single country, they are called “national models” [31–33].

3.1. Early Design Errors and Serious Difficulties Due to the Large Data and Program Collection Characteristics

In this Subsection, we present some serious difficulties/problems, inherent to large data and program collection characteristics. We will describe below some of the most severe difficulties of this type:

(a) A prevailing design mentality in connection with the first transport DS systems was the following: “Get all related or even loosely related data first and then organize them”. Practice, though, proved that this design mentality leads to disasters. In fact, it turned out that:
• This way of data acquisition does not guarantee a precise/consistent, methodological collection of data. On the contrary, the random, heuristic and ad hoc data collection, as a rule, results in a core with data and programs that are non-interoperable, internally inconsistent, practically impossible to organize or update properly and impossible to upgrade.
• Due to the aforementioned problems concerning the selected data, the core of the corresponding DSS practically collapses under the weight of the non-organizable data collections.

Consequently, this design methodology has been rather quickly abandoned.

(b) It must be pointed out that, after this early approach, the designers of a large-scale decision support system apply carefully designed data acquisition methodologies. However, it turns out that in practical cases, there are several characteristics, inherent to the elements of large data collections, which create severe difficulties and problems to the designers of the information system to be used as a platform of the DSS. Some of these characteristics will be presented immediately below.

(i) As a rule, in a very large DS system, the data are multi-thematic; hence, it is logical to expect a great variability of definitions, meanings and contents associated with the data.
(ii) The aforementioned variability is frequently accompanied by a great diversity of forms and formats of the elements of the data collection. This diversity practically always asks for a different handling of each subset of similar data.
(iii) Complexity is another serious problem associated with large-scale data collections and the corresponding DS systems. The term “complexity” is used in order to describe data with a particularly great number of interconnections, inter-relations and interdependencies.
(iv) Data polymorphism can easily destroy an arbitrary DS system. For example, the most common computer object in transport, the “transport link”, may literally have tenths or even hundreds of different definitions, meanings and data contents. Thus, the link between two cities A and B may refer to:

- Various road connections, where each road has its own characteristics (e.g., a highway with a given number of lanes with or without tolls, a national road, a secondary street, etc.);
- Various train connections, each one frequently with its own characteristics, such as high-velocity trains (TGV), intercity trains, local trains, commercial trains, etc.;
- Airplane connections, with airports having different connections with the city center;
- Seaways, where each way frequently includes ferries, container carriers, bulk carriers, tankers, cruisers, etc.;
- Inland navigation, etc.

In addition, the information concerning a single object in the DSS drastically depends on the purpose of the use of this object. This fact makes large-scale decision support system designers employ ad hoc solutions, which usually make upgrades and thematical shifts much more difficult or even impossible. An efficient solution to this problem will be presented later in Sections 5 and 6 of the present work.

(c) In many cases, large data and program collections must be distributed in several systems, which may be geographically dispersed. Achieving an efficient operation of each such sub-system, as well as maintaining a proper communication among the individual systems, while keeping the large data and program core healthy and consistent, is a difficult task indeed.

(d) The huge numbers of external data sources of large DS systems: In a considerable number of cases, the large DSS data core must handle data from a huge number of different, heterogeneous
and autonomous external sources, such as databases and/or various data collections. For example, the European Transport Information System for Policy Support (ETIS) receives data from more than 500 external sources [20], which are indeed heterogeneous and autonomous. Handling this number of different sources is a very difficult task and, in most cases, tackling this task is by no means automatic. Indeed, usually, a large initial effort by experts is required, in order to create a first version of an interoperable, internally consistent data and program core [34].

(e) In a large data collection, a considerable number of differences in the meaning and/or definition of common variables, coming from different sources, may appear. Thus, the task of defining the exact meaning and form of the variables that are common to many sources of the DS system may prove to be extremely difficult.

For example, in the first stages of the ETIS design, it has been proved that the difficulty of defining the term/variable “long distance trips” [35] was almost insurmountable. In fact, for the Scandinavians, this term referred to a distance greater than 100 km, while such an option practically excluded trips in the Netherlands entirely. In addition, other countries expressed a strong desire to associate with this term the age of the passengers and/or to incorporate in it whether the relevant trip is transit or not. Today, one may think that the term “long distance trips” must also include as a sub-variable the information whether the vehicle (automobile, motorcycle or train) is electric or conventional.

Thus, it not a surprise that a new, complete definition of the term “long distance trips” was given by the Transport General Directory and that Eurostat issued a directive to all EU Members to conform with this new definition and to always provide all relevant information.

(f) The next difficulty concerns the “level of detail” (abbreviated as LoD) or “granularity”. We shall try to clarify the content and importance of the term “level of detail” by means of the following example, concerning Google Maps. In fact, this application starts from offering a global earth projection. Then, the user may gradually increase the level of detail at an arbitrary point of the projection, according to his/her desire. At each such selection of the user, a more detailed map of the associated area appears on the screen. At certain levels, new geographical objects appear on the screen. For example, one may select a specific country and see its map. Subsequently, one may choose an area of this country and see the corresponding map in more detail. After a sequence of successive selections, where each one of them offers a greater level of detail, one may suddenly see some cities as points and some sketches of roads; hence, new objects appear. Further increase of the detail may offer more and more dense maps of the desired area. If one selects the option “satellite”, more and more detailed images of the selected area appear on the screen. In the higher detail level, the user may see individual public buildings and constructions, houses and details of them.

In the first steps of large DS system implementation, many designers considered that it is sufficient to include the higher level of detail only in the information system. However, the two aforementioned examples, but also many other DS systems that do not necessarily include geographical information, demonstrate that the approach “include the highest level of detail only” is completely erroneous. On the contrary, as it will be analytically presented in Sections 5 and 6 of the present work, a sound and systematic method for handling the levels of detail in a large DS system is to include all necessary information, associated with any level of detail, according to each issue and problem that the DSS has to tackle. In other words, the levels of detail in each application of the DSS strongly depend on the issue and problem in hand; therefore, the corresponding information system must be designed by keeping this issue dependence in mind.
3.2. Extreme Problems in Deploying, Managing, Maintaining, Updating and Upgrading Very Large-Scale Decision Support Systems

The deployment of a (very) large-scale decision support system meets with severe difficulties, due to the reasons described immediately below, which are in strong accordance with the content and the observations of [36,37]:

(a) Since the DSS platform is not sustainable in many cases [38], an effort for redesigning the DSS must be repeated after a short time period, when considerable changes must be accommodated. We would like to emphasize that the time period, after which it is necessary to redesign a certain DS system [39] (p. 4) [40,41], may frequently be as short as few years only, for example, 4 or 5 years.

(b) Serious difficulties in updating a large-scale decision support system: In such a system, the great number of the involved data, the significant complexity and polymorphism of them, the heterogeneity and dispersion of the data sources, as well as the fact that the various source data bases have been designed with a different mentality/approach, make the necessary, regular updating of the overall information system too difficult or even impossible. In fact, the problem is far more severe. The aforementioned factors render DSS upgrading much more difficult; as a rule, upgrading comprises improvement of existing data sets and programs, as well as incorporation of new data sets and programs, so that additional problems can be tackled by the DSS.

(c) Upgrading a large-scale decision support system is, very frequently, impossible: We would like to point out that, when a serious upgrading of a system is required, then, as a rule, a new DS system is frequently designed ad hoc, without a systematic methodology; this approach asks for a migration of the data and program core of the old system to the new one, a task which is extremely difficult and, in certain cases, impossible to achieve. Evidently, the new DS system will definitely manifest the very same problems as the old one, concerning updating and upgrading.

(d) The DSS lifecycle may strongly depend on the duration of the incumbency of the decision authorities: It is very well known that, in western-type democracies, the authorities who take decisions frequently change in a smooth way. This fact may render the time duration of the incumbency of the policy makers comparable to the time period, which is necessary for the deployment of a new or improved DS system. Thus, for example, very often, the policy priorities may drastically change, as a result of changes in the decision authorities. In this way, a DS system that has just begun to work may become partially or totally obsolete. Consequently, the DS system must be drastically changed or even redesigned from the beginning. Hence, it is imperative to design information systems in such a way so that maximum re-use of the existing data and programs can be achieved.

4. Contemporary, Mainstream Methodologies for Creating Large-Scale Decision Support Systems

First of all, a complete/thorough analysis and study, concerning the set of the required indicators, usually precedes the system design. This part of the DSS project is absolutely necessary; otherwise, corrections and additions at a later stage may be difficult or even impossible. The initial study must clearly and comprehensively offer information about the following, which will also prove very useful in the analysis of Sections 5 and 6:

(a) A comprehensive list of the general issues/problems, for which the DS system must provide support. We will use for it the term “thematic list”.

(b) A decomposition of each one of the aforementioned issues/problems to subcategories, further subcategories, etc., up to the fundamental, non-separable issues/problems. We will employ for this process the term “thematic decomposition”. For example, the general transport issue of “social/environmental impacts of transport” must initially discriminate between the subcategories “harmful impacts” and “beneficial impacts”. Then, the subcategory of “harmful impacts” must, among others, include:
• “Noise”;
• “Congestion”;
• “Pollutant emissions”.

For example, a complete associated path could be the following: transport policy issues → environmental impacts of transport → harmful environmental impacts of transport → noise → noise produced at the transport network → elementary/non-separable issue: noise produced by road traffic outside of the urban areas → indicator(s): noise measurements for all links and nodes of the road network outside of the urban areas.

On the other hand, the subcategory “beneficial impacts” must, among others, include:

• “Improved connectivity” (e.g., better island connectivity);
• “Transport time reduction”;
• “Growth” (e.g., residential or commercial relocation for a better environment, improvement of trade conditions, economic growth, etc.).

It must be emphasized that the form of an indicator can greatly vary, from a simple set of numbers (e.g., the former Kyoto indicators for environmental pollution) [42–45] to a complete database (e.g., a map indicating the road congestion of an area, with seasonal and daily indications).

(c) An ensemble of methods for either acquiring the indicators or computing them from the relevant data, using proper mathematical models.

(d) A complete list of the data, from which indicators must be evaluated.

(e) A complete list of the corresponding software, that is, programs that will implement the aforementioned abstract objects.

(f) A sufficient/adequate ensemble of “metadata”, which are necessary for the description of all indicators, data, acquisition or evaluation methods and the justification of their choice. It must be emphasized that this step is quite often partially considered or even neglected in the design of impressively numerous very large-scale information systems. Designing a very large-scale information system and including a poor/insufficient set of metadata is a grave error that will practically render the system non-maintainable, non-upgradable and non-modifiable.

The related information is, as a rule, a result of tedious efforts of experts, usually belonging to numerous scientific disciplines. Evidently, the work of the experts is supported by information technology and communication (ITS) engineers.

As a next step, ITS experts, using the aforementioned abstract results, implement the following:

• A user interface architectural tier, that is, an adequate number of computer systems to handle the communication of the users with the DSS system. Nowadays, this is achieved through the Internet and/or an Intranet, if confidentiality reasons dictate so.

• One or more databases and/or data warehouses specifically designed to handle large ensembles of data. The entire set of these tools is frequently called “data tier”. We note that the data tier may include specialized hardware, such as hardware “storages” or other similar components.

• A cluster of computer systems, sometimes called “application tier”, implementing the computational methods for the evaluation of the indicators, using data from the previous data tier. Additionally, the ITS engineers may include custom software for satisfying particular needs of the users of a DSS system for data visualization and/or specialized processing.

• A set of computer systems for the communication and mediation of the DSS with external data bases.

Contemporary solutions may include even more advanced software tools, such as “web services” and “containers” [46]. Moreover, in recent designs, several servers remain totally “in-memory”, in order to achieve highly increased performance; however, this solution is particularly expensive.
Next, suppose that a team of engineers implements a DS system ad hoc, without incorporating a rich set of metadata in it. In fact, without the complete knowledge of the appropriate metadata, the programs and data of an information system are practically entangled black boxes, to which corrections, maintenance, improvements, expansion and updating are difficult, if not impossible to be realized. In addition, a considerable number of users themselves are experts and/or they are supported by teams of experts; consequently, they will definitely need to know the details of acquisition and evaluation of the indicators. Thus, we would like to emphasize that the limiting factor in the design of large-scale decision support systems is not the hardware or software capabilities, but the very organization of the huge information itself.

5. A Novel, Systematic Methodology for Designing, Maintaining and Upgrading a Very Large-Scale Decision Support System

It has been previously stated that the design and implementation of a very-large scale decision support system is realized by means of the following two general stages:

Stage 1: A group of experts, after an extensive analysis of the problems/requirements in hand, ends up with an abstract structure, which determines all due operations of the DS system.

Stage 2: Based on this abstract structure, a group of ICT engineers produces an ad hoc implementation of the sought-for information system. However, as it has been more extensively described in Section 4 of the present work, any ad hoc implementation intrinsically suffers from serious flaws and problems of maintenance and upgrading of the developed system.

Hence, in the present section, a novel, systematic approach that circumvents these difficulties is described. This new approach includes the following steps:

(a) Development of a specialized meta-database, able to store the entire aforementioned abstract structure produced by the experts, using a corresponding template, which is more analytically described below. For this meta-database, we shall use the name “teleological meta-database” [47] or simply “meta-database”, for reasons that will become clear in the following steps.

(b) Insertion in this template of the thematic entities/entries of the system and the associated “thematic decomposition”, which ends up to the “non-separable objects” of the DSS in hand; these actions and terms will be more extensively analyzed in Section 5.1.

(c) Further filling of the template of the teleological meta-database with all the necessary indicator descriptions. At the same time, the analytic description of the exact method(s) of acquiring or evaluating these indicators must be included in this meta-database.

(d) Further inclusion in the teleological meta-database of the complete list of characteristics of the DSS external sources, together with all necessary information for accessing them.

(e) Incorporation of all the documentation metadata in the teleological meta-database; this documentation must fully cover the entire set of data, programs and other components and actions of the DSS.

We would like to emphasize that we render this teleological meta-database the “heart” of the entire DS system and its operations.

5.1. The Part of the Template that Includes the Complete Thematic Decomposition

As it has been already pointed out in Section 4, the first part of the template concerns the complete thematic list of the topics or issues for which the DSS must provide information/answers. An exhaustive thematic decomposition is applied to this list, expressed via a tree-like graph structure, down to further-non-separable objects/entities. We noticed that in the case of polymorphic data, the final leaves of the thematic decomposition usually highly depend on the sub-problem in hand. In other words, a specific object must be further decomposed, when the DSS tackles a certain problem, say A, while it may be a non-separable entity, when the system offers information concerning another problem, say B.
An abstract visualization of this tree-like, graph structure that performs such a thematic decomposition is depicted in Figure 1. In fact, concerning Figure 1, we would like to point out the following:

- The entire thematic list of the DSS incorporates issues $I_1, I_2, \ldots, I_N$.
- Each such issue is decomposed into successive sub-issues, for example, issue $I_1$ is at the first level decomposed into sub-issues $SI_{1,1}, SI_{1,2}, \ldots, SI_{1,K_1}$. Similarly, issue $I_N$ is decomposed into sub-issues $SI_{N,1}, SI_{N,2}, \ldots, SI_{N,K_N}$.
- In an analogous manner, sub-issue $SI_{1,1}$ is decomposed into $SI_{1,1,1}, SI_{1,1,2}, SI_{1,1,3}, \ldots, SI_{1,1,M_1}$, and so on.
- The decomposition process in connection with each initial thematic issue stops, when all non-separable objects, associated with this issue have been determined. In this way, for each issue $I_J$ ($J = 1, 2, \ldots, N$), a specific number of paths is obtained. We emphasize that the number of these paths as well as the depth of each path, as a rule, are not constant. We employ the symbol $NS$ for the non-separable object of each such path. Thus, for example, issue $I_1$ is eventually decomposed into $F_1$ non-separable objects, giving rise to "last-leaves" $NS_{1,1}, NS_{1,2}, \ldots, NS_{1,F_1}$.

![Thematic List of the System](image)

*Figure 1.* The structure of a thematic decomposition, an analytic explanation of which is presented in Section 5.1, where “MD” stands for “metadata.”
We repeat that, often, the length of the path that leads to NS \(_{1,1}\) may be different from the length that leads to NS \(_{1,2}\), which in turn may be different from the length of the path that leads to NS \(_{1,3}\), and so on. These differences in length are indicated in Figure 1 by the fact that the non-separable objects are placed at different levels.

Evidently, issue \(I_2\) ends up to the non-separable objects NS \(_{2,1}\), NS \(_{2,2}\), ..., NS \(_{2,F_2}\) and so forth. We stress that, in many instances, issue \(I_1\) may end up to an arbitrary number of non-separable objects NS \(_{1,R}\) following various different paths. In this case, we choose as length of the path that links \(I_1\) with NS \(_{1,R}\), the length of the path that connects these two entities, which starts at the higher-most level, namely, the one closer to \(I_1\), and so on.

- We have used a double arrow to indicate a connection of two arbitrary objects of the tree-like structure, in order to make clear that each sub-issue must “know” all previous sub-issues and the initial issue(s) that have generated it.
- We would like to emphasize that each issue and sub-issue is naturally linked to a set of metadata.
- The dotted arrows have been used to indicate that the entire structure is not a tree, but a graph, most probably, an acyclic one.
- The gray arrows that originate at the final non-separable objects of this tree-like, graph structure of Figure 1 manifest that this structure is linked to another one, that deals with the indicators and/or the data of the system.

5.2. The Template Part Covering the DSS Indicators and/or Data and Their Acquisition or Evaluation Method(s)

Each issue and/or each sub-issue, described in Section 5.1 and Figure 1, may be linked (and is indeed very frequently linked) to an ensemble of indicators, namely, certain quantitative characteristics of the issues and/or sub-issues in hand. Indicators help the user(s) of the DS system to make decisions as objectively as possible. As we have already pointed out, a simple, but characteristic, example of an important indicator is the noise generated by vehicles at various specific points of a road network [48]. Another crucial indicator is the volume of pollutants emitted by vehicles at these points [49].

We will use the term “the complete list of indicators” to express the entire set of indicators, incorporated in a DS information system.

Each indicator must be associated with (a) the data necessary for its computation and/or (b) the exact method(s) for acquiring or evaluating it.

As a rule, there are three ways of obtaining indicators, namely:

(a) For each indicator (or group of indicators), the DS system uses a specific computing methodology, which employs certain data in order to evaluate the indicator(s).
(b) The indicator (or group of indicators) may come directly from external sources, needing no further (or minimal) processing, and they are stored intact in the DS system.
(c) Finally, there are groups of indicators, which are the results of studies and/or development of external group of experts. These indicator(s) are also stored intact in the DS system.

An abstract representation of all three methods for acquiring or evaluating indicators is shown in Figure 2.

As far as this part of the template is concerned, one more important issue must be emphasized. A very consistent approach to describe a method that computes one or more indicators, so that the DS system remains maintainable and upgradable, is to decompose each such computational method into lesser “functional blocks” and describe all the necessary internal communication of these blocks (i.e., for each block, its input intermediate variables, the origins of these input variables, the output intermediate variables and the destination of these output variables must be described). This can be better understood by means of a simple, but fictitious, practical example, which is schematically shown in Figure 3. In this example, a certain computational method is decomposed into four (4) well-defined sub-blocks, namely, “model 1”, “model 2”, “sub-method 1” and “sub-method 2”.

\[ ... \]
Algorithms 2020, 13, x FOR PEER REVIEW 11 of 18

Figure 2. The structure of the meta-database template concerning the indicators, local data and computational methodologies.

Figure 3. An example of a first step of decomposition of a computational method. Any indicator and/or local datum may be compound, i.e., it may comprise a structure of simpler data.

Then, each sub-method (namely, model 1, …, sub-method 2) may be further decomposed in an analogous manner and so on. The decomposition stops when the experts that have initially developed the system feel that they have described the entire computational method adequately, so that another expert (or group of experts) can fully understand it at a later stage.

In Figure 4, another example is shown, which is actually encountered in the transport application area. In fact, the example refers to a specific method that computes the noise levels at a number of certain road network traffic links.
Figure 4. Computation of the levels of noise at certain links of a road network.

First, a traffic forecast model uses traffic counts on the links of the network and traffic growth factors (local data) to forecast traffic volumes for a future planning period. At the next step, a sub-computational method takes these forecasted traffic volumes and road capacity figures and computes the average vehicle velocities. Finally, a noise prediction model takes the computed average vehicle velocities, forecasted traffic volumes at the links of the network, as well as other road network characteristics and estimates the noise levels at these links of the network for the future planning period. The description of the computational methods starts at this level of abstraction. The sub-methods, shown in Figure 4, should be further decomposed. Once more, the decomposition stops when the experts that developed the system feel that the entire description is sufficient.

5.3. Including the External Sources and the Way to Access Them in the Template

We use the term “external sources” to describe the set of geographically dispersed databases linked to the main decision support system; these data bases are, as a rule, autonomous and most probably heterogeneous. Each such database may offer data and/or indicators to the main DS system. In turn, the main system may use these external data or indicators intact or employ them, in order to compute and/or generate other data or indicators. The part of the system that implements the communication of each external source with the main system and, possibly, computes all necessary quantities, is called “mediator”; the corresponding process itself is called “mediation”.

As mentioned before, each external source provides a part of the data of the main system and/or a part of the indicators. Consequently, the proposed meta-database must include:

(a) A complete description of the information furnished by the external sources.
(b) A full documentation concerning the method(s), which the mediator employs in order to evaluate the various indicators or data.
(c) The exact methodology and the frequency with which the data and/or indicators reaching the main system must be updated.
(d) Since a part (sometimes considerable) of the indicator ensemble may come from studies/R&D projects, these studies or projects can be considered as an “off-line” kind of external sources. It is necessary to describe the information associated with these indicators adequately, so that future expert users of the DS system can clearly and unambiguously understand the following:

- The quality of the indicators;
- Any possible special characteristics these indicators possess;
• Any limitations of the use of the group of specific indicators.

5.4. Including a Complete, Structured Description of Experts’ Knowledge in the Teleological Meta-Database

Figures 1 and 2 depict a fundamental characteristic of the meta-database, expressed by the fact that a “metadata box” is attached to every entity in the template. Each such box should incorporate the relevant knowledge of the experts who built the system and it should explain every choice of them (such as the thematic list, all issues, sub-issues, indicators, data, computational methods, external sources, etc.). The content of the “metadata box” should be written in such a way, so that at a later stage, experts, belonging to the same scientific discipline, can understand any special characteristic and any limitation of the use of all entities in the decision support system. The level of description must be detailed enough, so that these other experts can take over the maintenance and upgrade of the DS system. We must emphasize that confidentiality and/or intellectual property issues can be resolved by letting:

(a) The external sources control access to sensitive information, that is, information covered by GDPR; and/or, probably,
(b) The administration of the main core of the DSS have analogous rights.

Preferably, all this information must be further organized by building a specifically oriented sub-meta-database, which will provide all possible ways for accessing this information (e.g., accredited keywords, characteristic terms and methodology names). Additionally, a knowledge base [50,51] may be included in the DS system, with references to other related, important information.

5.5. Incorporating the Proper Documentation and Navigation Information for the Users in the Teleological Meta-Database

As it has been many times emphasized before, a large-scale decision support system is very complex. Thus, even a user expert in the related scientific field cannot use the system properly unless he/she is suitably advised. Therefore, it is rather imperative that specific documentation and navigation information must be included in the meta-database, so that a user, who is not familiar with the DS system, can benefit from it with minimal difficulty. There are two aspects of the aforementioned kind of information:

• The first is scientific: The expert designer of the system has to provide proper content titles or even content synopses, that can facilitate the unfamiliar user in navigating the DS system easily and without mistakes.

• The second one is technical: It aims at helping the user to understand the information system and the user interface. It must be provided by the ITC expert(s) who built the IS.

Overall, the documentation template of the meta-database can be built in such a manner, so as to assist all user experts and their collaborators to find out themselves how to use the DS system as easily as possible.

6. Substantial Advantages of a Decision Support System Developed on the Basis of the Proposed New Teleological Meta-Database

As it has been already discussed, the meta-database must have the following characteristics and properties:

(a) It must include the purpose and use of the employed indicators and/or of each datum.

(b) It must incorporate the entire thematic list, that is, all issues and sub-issues and the exact way each one of them is linked to the entities of (a) immediately above.

(c) All software programs of the DSS must be directly linked to the entities referred to in (a) and (b) immediately above.
(d) The aforementioned software programs of the DSS must be very well documented, and this
documentation must be properly placed in the template of the meta-database in a clear-cut manner.

Consequently, the proposed metadata structure can be considered as an exact and thorough
content description and functional map of the DS system in hand.

One main novelty of the methodology described here is that during the entire life cycle of the
DS system, a group of experts and engineers, even radically different than the one that developed
the specific information system, may achieve the following, by employing the proposed teleological
meta-database:

(a) Correct and efficient updating of the DS system.
(b) Make changes to the content of the DSS.
(c) Upgrade and/or expand the content of the DS system, by including further classes of issues,
    sub-issues, data, indicators, associated computational methods, new external sources, together
    with the corresponding mediators, etc. In fact, by exploiting the structure of the meta-database,
    one may achieve a consistent, considerable growth of the DS system. Without the use of the
    proposed teleological meta-database, there is a serious risk that such growth of the DSS may
    render it unmanageable.
(d) Reduce the content of the DS system, by eliminating a subset of the aforementioned entities,
    without causing any damage to the remaining system at all.

We would like to emphasize that any change imposed on the system must be accurately
reflected/incorporated in the meta-database template. In this way, the DSS will maintain all the
very good properties mentioned above; in the opposite case, all aforementioned actions concerning the
DS system will not be feasible.

We would like to point out that any part of the information of a DS system based on the
meta-database may be georeferenced/geographical. Actually, the use of the meta-database permits a
very advanced architecture of the georeferenced/geographical information, as it will be described in
another work.

7. Conclusions

Recent advances in information technology and communications allow for the development of
very large-scale decision support systems. However, the complexity, the polymorphism and the size
of the data and/or program core of such a very large DS system generate a number of very serious
problems, which are outlined here. These problems have to do with:

(a) The efficient updating of the DS system;
(b) The flexible expansion of the DSS;
(c) The consistent reduction of the information system;
(d) The ability to modify the system;
(e) The upgradability of the DSS, etc.

Moreover, it is argued that any ad hoc/non-strictly methodological development of a very
large-scale DS system will definitely lead to very serious intrinsic problems.

In the present manuscript, it was shown that only a consistent, fully methodological approach
can resolve all the aforementioned problems and it can drastically minimize the associated risks and
hazards. Such a methodological approach was explicitly introduced here. This methodology was
associated with the development of a very detailed, consistent and rigorous meta-database, which fully
describes the DS system and it includes:

(a) A complete thematic decomposition;
(b) A complete list of all indicators and/or data and their acquisition or evaluation method(s);
(c) A complete list of all the external sources of the system, together with the way to access them;
(d) An absolutely sufficient and structured description of the experts’ knowledge concerning the whole system;
(e) The proper documentation and navigation information for users in this (teleological) meta-database.

Author Contributions: Conceptualization, E.K., C.P. and A.B.; Investigation, G.T., N.V.K. and A.B.; Project administration, E.K.; Resources, G.T., N.V.K. and E.M.; Supervision, E.K., C.P., G.T., N.V.K., A.B. and E.M.; Writing—original draft, E.K., C.P. and G.T.; Writing—review & editing, G.T., N.V.K., A.B., E.M. and A.R.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Fagerholt, K. A computer-based decision support system for vessel fleet scheduling—Experience and future research. Decis. Support Syst. 2004, 37, 35–47. [CrossRef]
2. Fagerholt, K.; Johnsen, T.A.V.; Lindstad, H. Fleet deployment in liner shipping: A case study. Marit. Policy Manag. 2009, 36, 397–409. [CrossRef]
3. Clemente, M.; Fanti, M.P.; Iacobellis, G.; Nolich, M.; Ukovitch, W. A Decision Support System for User-Based Vehicle Relocation in Car Sharing Systems. IEEE Trans. Syst. Man Cybern. Syst. 2017, 48, 1283–1296. [CrossRef]
4. Bolman, B.; Jak, R.G.; Van Hoof, L. Unravelling the myth—The use of Decisions Support Systems in marine management. Mar. Policy 2018, 87, 241–249. [CrossRef]
5. Gil, M.; Wróbel, K.; Montewka, J.; Goerlandt, F. A bibliometric analysis and systematic review of shipboard Decision Support Systems for accident prevention. Saf. Sci. 2020, 128, 104717. [CrossRef]
6. Pieri, G.; Cocco, M.; Salvetti, O. A Marine Information System for Environmental Monitoring: ARGO-MIS. J. Mar. Sci. Eng. 2018, 6, 15. [CrossRef]
7. Caris, A.; Macharis, C.; Janssens, G.K. Decision support in intermodal transport: A new research agenda. Comput. Ind. 2013, 64, 105–112. [CrossRef]
8. Qaiser, F.H.; Ahmed, K.; Sykora, M.; Choudhary, A.K.; Simpson, M. Decision support systems for sustainable logistics: A review and bibliometric analysis. Ind. Manag. Data. Syst. 2017, 117, 1376–1388. [CrossRef]
9. Dotoli, M.; Epicoco, N.; Falagario, M.; Seatzu, C.; Turchiano, B. A Decision Support System for Optimizing Operations at Intermodal Railroad Terminals. IEEE Trans. Syst. Man Cybern. Syst. 2016, 47, 487–501. [CrossRef]
10. Zhan, J.; Ling, A.; An, N.; Li, L.; Sha, Y.; Li, X.; Liu, G. Building a Practical Ontology for Emergency Response Systems. In Proceedings of the IEEE Computer Society: Proceedings on, International Conference in Computer Science and Software Engineering CSSE ’08 2008, Hubei, China, 12–14 December 2008; Volume 4, pp. 222–225. [CrossRef]
11. Comfort, L.K.; Sungu, Y.; Johnson, D.; Dunn, M. Complex Systems in Crisis: Anticipation and Resilience in Dynamic Environments. J. Contingencies Crisis Manag. 2001, 9, 144–158. [CrossRef]
12. Sakellarious, S.; Tampakis, S.; Samara, F.; Stouragiris, A.; Christopoulos, O. Review of state-of-the-art decision support systems (DSSs) for prevention and suppression of forest fires. J. For. Res. 2017, 28, 1107–1117. [CrossRef]
13. Rauner, M.S.; Niessner, H.; Odd, S.; Pope, A.; Neville, K.; O’Riordan, S.; Sasse, L.; Tomic, K. An advanced decision support system for European disaster management: The feature of the skills taxonomy. Central Eur. J. Oper. Res. 2018, 26, 485–530. [CrossRef]
14. McCown, R.L. Changing systems for supporting farmers’ decisions: Problems, paradigms, and prospects. Agric. Syst. 2002, 74, 179–220. [CrossRef]
15. Salewicz, K.A.; Nakayama, M. Development of a web-based decision support system (DSS) for managing large international rivers. Glob. Environ. Chang. 2004, 14, 25–37. [CrossRef]
16. Ayed, M.B.; Lifi, H.; Kolski, C.; Alimi, A.M. A user-centered approach for the design and implementation of KDD-based DSS: A case study in the healthcare domain. Decis. Support Syst. 2010, 50, 64–78. [CrossRef]
17. Sisk, G.M.; Miles, J.C.; Moore, C.J. Designer Centered Development of GA-Based DSS for Conceptual Design of Buildings. *J. Comput. Civ. Eng.* 2003, 17, 159–166. [CrossRef]

18. Tavasszy, L.A.; Smeenk, B.; Ruijgrok, C.J. A DSS for modelling logistic chains in freight transport policy analysis. *Int. Trans. Op. Res.* 1998, 5, 447–459. [CrossRef]

19. Ocali-Akunal, E.V. Decision support systems in transport planning. *Procedia Eng.* 2016, 161, 1119–1126. [CrossRef]

20. Ballis, A. Implementing the European Transport information System. *Transp. Res. Rec. J. Transp. Res. Board* 2006, 1957, 23–31. [CrossRef]

21. Yatskiv, I.; Yurshevich, E. Data Actualization Using Regression Models in Decision Support System for Urban Transport Planning. In Proceedings of the 10th International Conference on Dependability and Complex Systems DepCoS-RELCOMEX 2015, Brunów, Poland, 29 June–3 July 2015; Volume 365. [CrossRef]

22. Damart, S.; Roy, B. The uses of cost–benefit analysis in public transportation decision-making in France. *Transp. Policy* 2009, 16, 200–212. [CrossRef]

23. Barfod, M.B.; Salling, K.B.; Leleur, S. Composite decision support by combining cost-benefit and multi-criteria decision analysis. *Decis. Support Syst.* 2011, 51, 167–175. [CrossRef]

24. Sacchelli, S. A Decision Support System for trade-off analysis and dynamic evaluation of forest ecosystem services. *iFOR-Biogosci.* For. 2018, 11, 171–180. [CrossRef]

25. Loseto, G.; Scioscia, F.; Ruta, M.; Gramagna, F.; Ieva, S.; Pinto, A.; Scioscia, C. Knowledge-Based Decision Support in Healthcare via Near Field Communication. *Sensors* 2020, 20, 4923. [CrossRef] [PubMed]

26. Spoladore, D.; Sacco, M. Semantic and Dweller-Based Decision Support System for the Reconfiguration of Domestic Environments: RecAAL. *Electronics* 2018, 7, 179. [CrossRef]

27. Moreira, M.W.L.; Rodrigues, J.J.P.; Korotaev, V.; Al-Muhtadi, J.; Kumar, N. A Comprehensive Review on Smart Decision Support Systems for Health Care. *IEEE Syst. J.* 2019, 13, 3536–3545. [CrossRef]

28. Encyclopaedia Britannica. Information System. Available online: https://www.britannica.com/topic/information-system (accessed on 6 December 2020).

29. Archive of the European Integration (AEI); European Sources Online (ESO). Information Guide: Trans-European Networks. University of Pittsburgh, Cardiff EDC, 2013. Available online: http://aei.pitt.edu/75428/ (accessed on 25 October 2020).

30. Port of Rotterdam. Port Community System. Available online: https://www.portofrotterdam.com/en/doing-business/services/service-range/port-community-system (accessed on 25 October 2020).

31. Gerard, J.; Gunn, H.; Walker, W. National and International Freight Transport Models: An Overview and Ideas for Future Development. *Transp. Rev.* 2004, 24, 103–124. [CrossRef]

32. De Jong, G.; Vierth, I.; Tavasszy, L.; Ben-Akiva, M. Recent developments in national and international freight transport models within Europe. *Transportation* 2013, 40, 347–371. [CrossRef]

33. Rich, J.; Nielsen, O.A.; Brems, C.; Hansen, C.O. Overall design of the Danish national transport model. In *Proceedings of the Annual Transport Conference*; Aalborg University: Aalborg, Denmark, 2010; Volume 17, ISSN 1603-9696.

34. Chiehyeon, L.; Kim, K.J.; Maglio, P. Smart Cities with Big Data: Reference models, challenges, and considerations. *Sci. Direct* 2018, 82, 86–99. [CrossRef]

35. Gerike, R.; Schulz, A. Workshop Synthesis: Surveys on long-distance travel and other rare events. *Transp. Res. Procedia* 2018, 32, 535–541. [CrossRef]

36. Sillitto, H.G. Design principles for Ultra-Large-Scale (ULS) Systems. *INCOSE Int. Symp.* 2010, 20, 63–82. [CrossRef]

37. Gabriel, R.P.; Northrop, L.; Schmidt, D.C.; Sullivan, K. Ultra-large-scale systems. In *Proceedings of the Companion to the 21st ACM SIGPLAN Symposium on Object-Oriented Programming Systems, Languages, and Applications (OOPSLA 2006)*, Portland, OR, USA, 22–26 October 2006; Association for Computing Machinery: New York, NY, USA, 2006; pp. 632–634. [CrossRef]

38. Arnott, D.; Dodson, G. *Decision Support Systems Failure*; Springer: Berlin/Heidelberg, Germany, 2008; Volume 1, pp. 763–790. [CrossRef]

39. Jovic, M.; Tijan, E.; Zgaljic, D.; Karanikic, P. SWOT analysis of selected digital technologies in transport economics. In *Proceedings of the 43rd International Convention on Information, Communication and Electronic Technology, Digitalizing Logistics processes (DIGLOGS)*, Opatija, Croatia, 28 September–2 October 2020.
40. Limitations & Disadvantages of Decision Support Systems. Available online: https://www.managementstudyguide.com/limitations-and-disadvantages-of-decision-support-systems.htm (accessed on 25 October 2020).

41. Disadvantages of Decision Support Systems. Available online: http://dsssystem.blogspot.com/2010/01/disadvantages-of-decision-support.html (accessed on 25 October 2020).

42. United Nations, Treaty Collection, Depositary, Environment. A Kyoto Protocol to the United Nations Framework Convention on Climate Change. Available online: https://treaties.un.org/pages/ViewDetails.aspx?src=TREATY&mtdsg_no=XXVII-7a&chapter=27&lang=en (accessed on 25 October 2020).

43. United Nations, Climate Change, Process and Meetings. The Kyoto Protocol-Status of Ratification. Available online: https://unfccc.int/process/the-kyoto-protocol/status-of-ratification (accessed on 25 October 2020).

44. Grubb, M.; Vrolijk, C.; Brack, D. Routledge Revivals: Kyoto Protocol (1999): A Guide and Assessment; Routledge: Abingdon-on-Thames, UK, 2018; ISBN 9781315147024. [CrossRef]

45. Oberthür, S.; Hermann, E.O. The Kyoto Protocol: International Climate Policy for the 21st Century; Springer Science & Business Media: Berlin, Germany, 1999. [CrossRef]

46. Dearle, A. Software Deployment, Past, Present and Future; Future of Software Engineering (FOSE): Minneapolis, MN, USA, 2007; pp. 269–284. [CrossRef]

47. Kavvadia, H. Teleology in Systems Analysis and Design Methods. SSRN Electron. J. 2020. [CrossRef]

48. Méline, J.; Van Hulst, A.; Thomas, F.; Chaix, B. Road, rail, and air transportation noise in residential and workplace neighborhoods and blood pressure (RECORD Study). Noise Health 2015, 17, 308–319. [CrossRef] [PubMed]

49. Chen, L.; Yang, H. Managing congestion and emissions in road networks with tolls and rebates. Transp. Res. Part B Methodol. 2012, 46, 933–948. [CrossRef]

50. Atymtayeva, L.; Kozhakhmet, K.; Bortsova, G. Building a knowledge base for expert system in information security. Soft Comput. Artif. Intell. Adv. Intell. Syst. Comput. 2014, 270, 57–76. [CrossRef]

51. Breuil, D. On knowledge based management systems: Integrating artificial intelligence and database technologies. Eur. J. Oper. Res. 1988, 33, 354–355. [CrossRef]

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).