



- (51) **International Patent Classification:**  
G06N 5/02 (2006.01)
- (21) **International Application Number:**  
PCT/MY2014/000149
- (22) **International Filing Date:**  
3 June 2014 (03.06.2014)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**  
PI 2013004362 3 December 2013 (03.12.2013) MY
- (71) **Applicant:** MIMOS BERHAD [MY/MY]; Technology Park Malaysia, 57000 Kuala Lumpur (MY).
- (72) **Inventors:** CHEW, Yew Choong; c/o Mimos Berhad, Technology Park Malaysia, 57000 Kuala Lumpur (MY). KOH, May Fern; c/o Mimos Berhad, Technology Park Malaysia, 57000 Kuala Lumpur (MY). A/L M. PERUMAL, Nagendran; c/o Mimos Berhad, Technology Park Malaysia, 57000 Kuala Lumpur (MY). TEH, Kian Lai; c/o Mimos Berhad, Technology Park Malaysia, 57000 Kuala Lumpur (MY). RAHMAT, Rafizah Binti; c/o Mimos Berhad, Technology Park Malaysia, 57000 Kuala Lumpur (MY).
- (74) **Agent:** MIRANDAH, Patrick; Mirandah Asia (Malaysia) sdn bhd, Suite 3B-19-3, Plaza Sentral, Jalan Stesen Sentral 5, 50470 Kuala Lumpur (MY).
- (81) **Designated States** (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (84) **Designated States** (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

[Continued on next page]

## (54) Title: A SYSTEM AND METHOD FOR DATA INTEGRATION

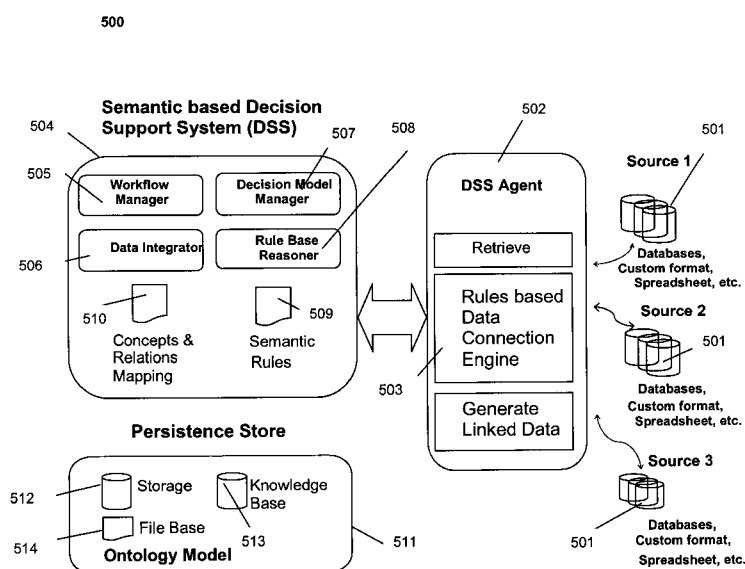


FIGURE 5

(57) **Abstract:** Data integration from various and variety of data sources is provided without changing the source of the data itself by connecting synonyms data based on predefined semantic rules and maintaining data originality. The system of the present invention comprising at least one decision support system (DSS) agent (502) comprising at least one rules based data connection engine (503) adapted to generate linked data from a plurality of data retrieved from a plurality of heterogeneous data sources; and at least one semantic based decision support system (DSS) (504) in communication with said decision support system agent and comprising a predefined set of semantic rules; and concepts and relations mapping information. Originality of the data will be maintained as none of the data source will be edited or removed, as compared to traditional algorithms based approach. The methodology of the present invention comprises steps of retrieving a plurality of data from a plurality of heterogeneous data sources (802); pre-defining set of semantic rules (804); forwarding predefined semantic rules into Rules Based Data Connection Engine to link schema-to-schema, property-to-property and schema to property of said data sources and to map schema and data values from said data sources based

on a set of predefined semantic rules (803); generating connected data from said mapped schema and data values (805); and returning relationships between schemas and values to user (806). Accuracy of intermediate and final results as obtained through the present invention is independent from the sequence of algorithms execution.

**Declarations under Rule 4.17:**

- *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*
- *of inventorship (Rule 4.17(iv))*

**Published:**

- *with international search report (Art. 21(3))*
- *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

## A SYSTEM AND METHOD FOR DATA INTEGRATION

### FIELD OF INVENTION

5 The present invention relates to a system and method for data integration. In particular, the invention relates to a system and method for data integration that involves connection of data from various data sources. Originality of data is maintained as synonyms data are connected based on pre-defined semantic rules utilizing Rules-based Data Connection Engine.

10

### BACKGROUND ART

Traditional algorithm based data cleansing processes usually involve the detection and correction of corrupt or inaccurate records from data sources. During the cleansing  
15 process, so called dirty data is replaced, modified or deleted. This process produces a smaller dataset in which some of the original information collected is permanent lost.

Data cleansing processes usually involve more than one algorithm. The selection of algorithms to be used and the execution sequence of the selected algorithms are always  
20 questionable issues as most algorithms selected are not symmetric to one another. Often, the algorithms are not dependent on one another, nor do they satisfy the Commutative ( $a + b = b + a$ ) and Associative ( $(a \times b) \times c = a \times (b \times c)$ ) laws.

An example of a traditional data cleansing process (100) is illustrated in Figures 1-4. As  
25 illustrated, data source (101) is collected from a variety of sources (102) and cleansed through a plurality of algorithms (103). From this process, a final result (104) is obtained. As shown in Figures 2 and 3 in slightly more detail, after application of a first algorithm (103a), and a first intermediate result (105a) is obtained. This is then passed through a second algorithm (103b) to obtain a second intermediate result (105b). This process is  
30 repeated through any number of algorithms (103c....103n), returning further intermediate results (105c....105n), until the final result (104) is reached.

The difficulty with this process lies in the selection of the algorithms (103) and determination of the best execution sequence of the algorithms (103). Each intermediate result (105) will invariably lose some of the original data source (101) input to the process (100). This is illustrated in Figure 4, in which data is collected from three data sources (102a, 102b and 102c). The data source (102a) is processed using a traditional cleansing process (100) to arrive at a final result 104. In this example, data source 1 (102a) provides for "Sex" with options "M" for "male" and "F" for "female". Data source 2 (102b) provides for "Jantina", the Malay word for "sex", with options "L" and "P", representing "male" and "female" respectively, rather than "M" and "F". Data source 3 (102c) provides for "Gender" with options "1" and "2". As such, the data collected may be combined if appropriate rules are employed. However, in traditional cleansing processes (100), the original schema and information may be lost in the final result (104). In this example, "Sex", "L", "P", "Jantina", "1" and "2" are lost in the process (100). That is, 2/3 of the original schema and information is lost.

United States Patent No. 8,166,048 describes a system for maintaining master reference data for entities. The system includes multiple reference data sets at multiple different data storages. A reference data set for a particular entity uniquely identifies the particular entity. The system also includes a first master reference data set at a first data storage that is at least as reliable as a second reference data set at a second data storage. In some embodiments, the first data storage can be updated through real-time process or an offline process (e.g. a batch process). The first master reference data set includes at least one data record, and content metadata regarding the data record. In some embodiments, content metadata comprises lineage data that includes each preceding value that was contained in the data record. In some of these embodiments, lineage includes other factors that affected the present and previous values contained in the data record.

United States Patent Publication No. 2012/0095973 A1 describes a method and system for developing data integration applications with reusable semantic types to represent and process application data. Methods include creating schemas to describe external data, creating semantic types to describe internal data, mapping schemas to semantic types, developing dataflow that configure input and output operations using schemas,

mappings, and semantic types and all other transformation operations and functions based solely on semantic types, and executing dataflow defined in this manner.

5 The present invention advantageously provides systems and methods that perform data integration by connecting data from a variety of data sources without changing the data source. Advantageously, data originality is maintained as none of the data source is edited or removed, which is usually the case using traditional algorithm based approaches. As such, accuracy of the intermediate and final results is advantageously independent of the sequence of algorithm execution.

10

The subject matter claimed herein is not limited to embodiments that solve any disadvantages or that operate only in environments such as those described above. Rather, this background is only provided to illustrate one exemplary technology area where some embodiments described herein may be practice.

## SUMMARY OF INVENTION

The present invention relates to a system and method for data integration. In particular, the invention relates to a system and method for data integration that involves  
5 connection of data from various data sources by maintaining data originality as synonyms data are connected based on pre-defined semantic rules utilizing Rules-based Data Connection Engine.

One aspect of the present invention provides a system (500) for data integration by  
10 connecting synonyms data based on pre-defined semantic rules and maintaining data originality. The system comprising at least one decision support system (DSS) agent (502) comprising at least one rules based data connection engine (503) adapted to generate linked data from a plurality of data retrieved from a plurality of heterogeneous data sources (501); and at least one semantic based decision support system (DSS)  
15 (504) in communication with said decision support system agent (502) and comprising a predefined set of semantic rules; and concepts and relations mapping information. The at least one rule based data connection engine (503) further having means for defining coverage of subject domain; defining schemas, properties and relationships; reusing schemas and properties; implanting domain coverage into relationships;  
20 ensuring relationships between schemas and values are coherent; mapping to knowledge base to increase matching; linking schema-to-schema, property-to-property and schema-to property incrementally; and returning connected schema and properties.

Another aspect of the invention provides that the said semantic based decision support  
25 system (DSS) (504) further comprises at least one workflow manager module (505); at least one data integration module (506), at least one decision model manager module (507); and at least one rule based reasoned module (508) adapted to link concepts and relations mapping and semantic rules into the rules based data connection engine (503).

30 A further aspect of the invention provides that the said semantic based decision support system (DSS) (504) is in communication with at least one persistence store (511) comprising at least one storage (512), at least one knowledge base (513) and at least one file base (504). The said persistence store (511) is an ontology model.

Another aspect of the invention provides a method (800) for data integration by connecting synonyms data based on pre-defined semantic rules and maintaining data originality. The said method comprising steps of pre-defining set of semantic rules (804); retrieving a plurality of data from a plurality of heterogeneous data sources (802);  
5 forwarding predefined semantic rules into Rules Based Data Connection Engine to link schema-to-schema, property-to-property and schema-to property of said data sources and to map schema and data values from said data sources based on a set of predefined semantic rules (803); and generating connected data from said mapped schema and data values (805); and returning relationships between schemas and values  
10 to user (806). The step of forwarding predefined semantic rules into Rules Based Data Connection Engine to link schema-to-schema, property-to- property and schema-to property of said data sources and to map schema and data values from said data sources based on a set of predefined semantic rules (803) further comprising steps of defining coverage of subject domain (901); defining schemas, properties and  
15 relationships (902); reusing schemas and properties (903); implanting domain coverage into relationships (904); ensuring relationships between schemas and values are coherent (905); mapping to knowledge base to increase matching (906); linking schema-to-schema, property-to-property and schema-to property incrementally (908); and returning connected schema and properties (909). The said semantic rules are a set of  
20 rules adapted to connect data from heterogeneous data sources.

Yet another aspect of the invention relates to a method additionally comprising establishing relationships between the heterogeneous data sources.

25 The present invention consists of features and a combination of parts hereinafter fully described and illustrated in the accompanying drawings, it being understood that various changes in the details may be made without departing from the scope of the invention or sacrificing any of the advantages of the present invention.

**BRIEF DESCRIPTION OF ACCOMPANYING DRAWINGS**

To further clarify various aspects of some embodiments of the present invention, a more particular description of the invention will be rendered by references to specific  
5       embodiments thereof, which are illustrated in the appended drawings. It is appreciated that these drawings depict only typical embodiments of the invention and are therefore not to be considered limiting of its scope. The invention will be described and explained with additional specificity and detail through the accompanying drawings in which:

10       FIG. 1 illustrates a generic flow diagram for a traditional data cleansing process.

FIG. 2 illustrates another flow diagram for a traditional data cleansing process.

FIG. 3 illustrates a further flow diagram for a traditional data cleansing process.

15

FIG. 4 illustrates an example of data collection and loss using a traditional data cleansing process.

FIG. 5 illustrates a system of an embodiment of the invention.

20

FIG. 6 illustrates an embodiment of the rules based data connecting approach of the invention.

FIG. 7 illustrates an example of rule based data connection in accordance with an  
25       embodiment of the invention.

FIG. 8 illustrates a simplified flow diagram of the methodology of an embodiment of the invention.

30       FIG. 9 illustrates a flow diagram of the operation of the rules based data connection engine.

## DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention relates to a system and method for data integration. In particular, the invention relates to a system and method for data integration that involves  
5 connection of data from various data sources by maintaining data originality as synonyms data are connected based on pre-defined semantic rules utilizing Rules-based Data Connection Engine.

Hereinafter, this specification will describe the present invention according to the  
10 preferred embodiments. It is to be understood that limiting the description to the preferred embodiments of the invention is merely to facilitate discussion of the present invention and it is envisioned without departing from the scope of the appended claims.

Referring to Figure 5, a system (500) for data integration by connecting synonyms data  
15 based on pre-defined semantic rules wherein originality of data is maintained is illustrated. According to this system (500) data source is retrieved from a plurality of heterogeneous data sources (501) through a decision support system agent (DSS Agent) (502). The DSS Agent (502) comprises a rules based data connection engine (503) that is adapted to receive the collected data and generate linked data from the  
20 data retrieved from heterogeneous data sources (501). The rule based data connection engine (503) having means for defining coverage of subject domain; defining schemas, properties and relationships; reusing schemas and properties; implanting domain coverage into relationships; ensuring relationships between schemas and values are coherent; mapping to knowledge base to increase matching; implanting domain  
25 coverage into relationships; linking schema-to-schema, property-to-property and schema-to property incrementally; and returning connected schema and properties.

The DSS Agent (502) is in communication with a semantic based decision support system (DSS) (504). The DSS (504) includes a workflow manager module (505), a data  
30 integration module (506), a decision model manager module (507) and a rule based reasoned module (508). Importantly, the DSS (504) is provided with a predefined set of semantic rules (509) and concepts and relations mapping information (510) that facilitate linking of schema-to-schema, property-to-property and schema-to property of the data

sources (501) and mapping of the schema and data values from the data sources (501) based on the set of predefined semantic rules (509). The said semantic based decision support system (DSS) (504) is adapted to link concepts and relations mapping (510) and semantic rules (509) into the rules based data connection engine (503). A persistence  
5 store (511) is also provided comprising storage (512), knowledge base (513) and file base (514). The said persistence store (511) is an ontology model.

A general method (600) of an embodiment of the invention is illustrated generically in Figure 6. Referring to that figure, again a plurality of heterogeneous data sources (601)  
10 generate data that is retrieved and processed by a rules based data connection engine (603) based on semantic rules and concepts and relations mapping provided in a DSS (604). As multiple rules are applied in a single process, issues faced by the application of multiple algorithms are advantageously avoided as data source is connected (605) to generate the final result (606).

15

An example of rules based data connection (700) is provided in Figure 7. For convenience, the example corresponds with that provided in Figure 4 relating to traditional algorithm based processing. Referring to Figure 7, data source (701) is retrieved from a plurality of heterogeneous data sources (702). Semantic rules (703) as  
20 provided in a DSS (704) are used to connect (705) the data source (701). In this way, data is not lost, as is the case with traditional processing as illustrated in Figure 4.

The methodology (800) of the present invention for data integration by connecting synonyms data based on pre-defined semantic rules and maintaining data originality is  
25 illustrated in Figure 8. As illustrated, a plurality of heterogeneous data sources (801) generates data and the set of semantic rules are pre-defined (804). Thereafter, the plurality of data is retrieved from a plurality of heterogeneous data sources (802) and processed in a rules based data connection engine (803). The set of semantic rules are used to process the data and generate connected data from mapped schema and data  
30 values (805). The said semantic rules are a set of rules adapted to connect data from heterogeneous data sources. The predefined semantic rules are forwarded into the rules based data connection engine to link schema-to-schema, property-to-property and schema-to property of said data sources and to map schema and data values from said

data sources based on a set of predefined semantic rules (803). This returns the final result wherein relationships between schemas and values are returned to the user (806).

Referring to Figure 9, the work flow (900) of the rules based data connection engine is illustrated. The process generally includes initially defining coverage of a subject domain (901) and defining schemas, properties and relationships (902). Schemas and properties are reused as much as possible (903). Implanting domain coverage into relationships (904) follows, and relationships between schemas and values are confirmed and ensured as coherent (905). Mapping to knowledge base is carried out to increase the likelihood of matching (906). Linking of schema-to-schema, property-to-property and schema-to property incrementally (908) is carried out and connected schema and properties returned. The said method further includes establishing relationships between said heterogeneous data sources. This generally completes the work flow (900).

The present invention is directed to integrate data from various and variety data sources, without changing the source of the data itself. Originality of the data will be maintained as none of the data source will be edited or removed, as compared to traditional algorithms based approach. Further, accuracy of intermediate and final results as obtained through the present invention is independent from the sequence of algorithms execution. Data is integrated by connecting synonyms data based on pre-defined semantic rules by utilizing rules-based data connection engine.

Unless the context requires otherwise or specifically stated to the contrary, integers, steps or elements of the invention recited herein as singular integers, steps or elements clearly encompass both singular and plural forms of the recited integers, steps or elements.

Throughout this specification, unless the context requires otherwise, the word "comprise", or variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated step or element or integer or group of steps or elements or integers, but not the exclusion of any other step or element or integer or group of steps, elements or integers. Thus, in the context of this specification, the term "comprising" is

- 10 -

used in an inclusive sense and thus should be understood as meaning "including principally, but not necessarily solely".

5 It will be appreciated that the foregoing description has been given by way of illustrative example of the invention and that all such modifications and variations thereto as would be apparent to persons of skill in the art are deemed to fall within the broad scope and ambit of the invention as herein set forth.

**CLAIMS**

1. A system (500) for data integration by connecting synonyms data based on pre-defined semantic rules and maintaining data originality comprising:

at least one decision support system (DSS) agent (502) comprising at least one rules based data connection engine (503) adapted to generate linked data from a plurality of data retrieved from a plurality of heterogeneous data sources (501); and

at least one semantic based decision support system (DSS) (504) in communication with said decision support system agent (502) and comprising a predefined set of semantic rules, and concepts and relations mapping information

characterized in that

the at least one rule based data connection engine (503) further having means for:

defining coverage of subject domain;

defining schemas, properties and relationships;

reusing schemas and properties;

implanting domain coverage into relationships;

ensuring relationships between schemas and values

are coherent;

mapping to knowledge base to increase matching;

linking schema-to-schema, property-to-property and

schema-to property incrementally; and

returning connected schema and properties.

2. A system (500) according to claim 1, wherein said semantic based decision support system (DSS) (504) further comprises at least one workflow manager module (505); at least one data integration module (506), at least one decision model manager module (507); and at least one rule based reasoned module (508) is adapted to link concepts and relations mapping (510) and semantic rules into the rules based data connection engine (503).

3. A system (500) according to claim 1, wherein said semantic based decision support system (DSS) (504) is in communication with at least one persistence store (511) comprising at least one storage (512), at least one knowledge base (513) and at least one file base (504).

5

4. A system (500) according to claim 1, wherein said persistence store (511) is an ontology model.

5. A method (800) for data integration by connecting synonyms data based on pre-defined semantic rules and maintaining data originality comprising steps of:

10

pre-defining set of semantic rules (804)

retrieving a plurality of data from a plurality of heterogeneous data sources (802);

forwarding predefined semantic rules into Rules Based Data

15

Connection Engine to link schema-to-schema, property-to-property and schema-to property of said data sources and to map schema and data values from said data sources based on a set of predefined semantic rules (803);

generating connected data from said mapped schema and data values (805); and

20

returning relationships between schemas and values to user (806) characterized in that

forwarding predefined semantic rules into Rules Based Data

Connection Engine to link schema-to-schema, property-to-

25

property and schema-to property of said data sources and to map schema and data values from said data sources based on a set of predefined semantic rules (803) further

comprising steps of:

defining coverage of subject domain (901);

30

defining schemas, properties and relationships (902);

reusing schemas and properties (903);

implanting domain coverage into relationships (904);

- 13 -

ensuring relationships between schemas and values  
are coherent (905);  
mapping to knowledge base to increase matching  
(906);  
5 linking schema-to-schema, property-to-property and  
schema-to property incrementally (908); and  
returning connected schema and properties (909).

6. A method according to claim 5, wherein said semantic rules are a set of rules  
10 adapted to connect data from heterogeneous data sources.
7. A method according to claim 5, further comprising establishing relationships  
between said heterogeneous data sources.

1/9

100

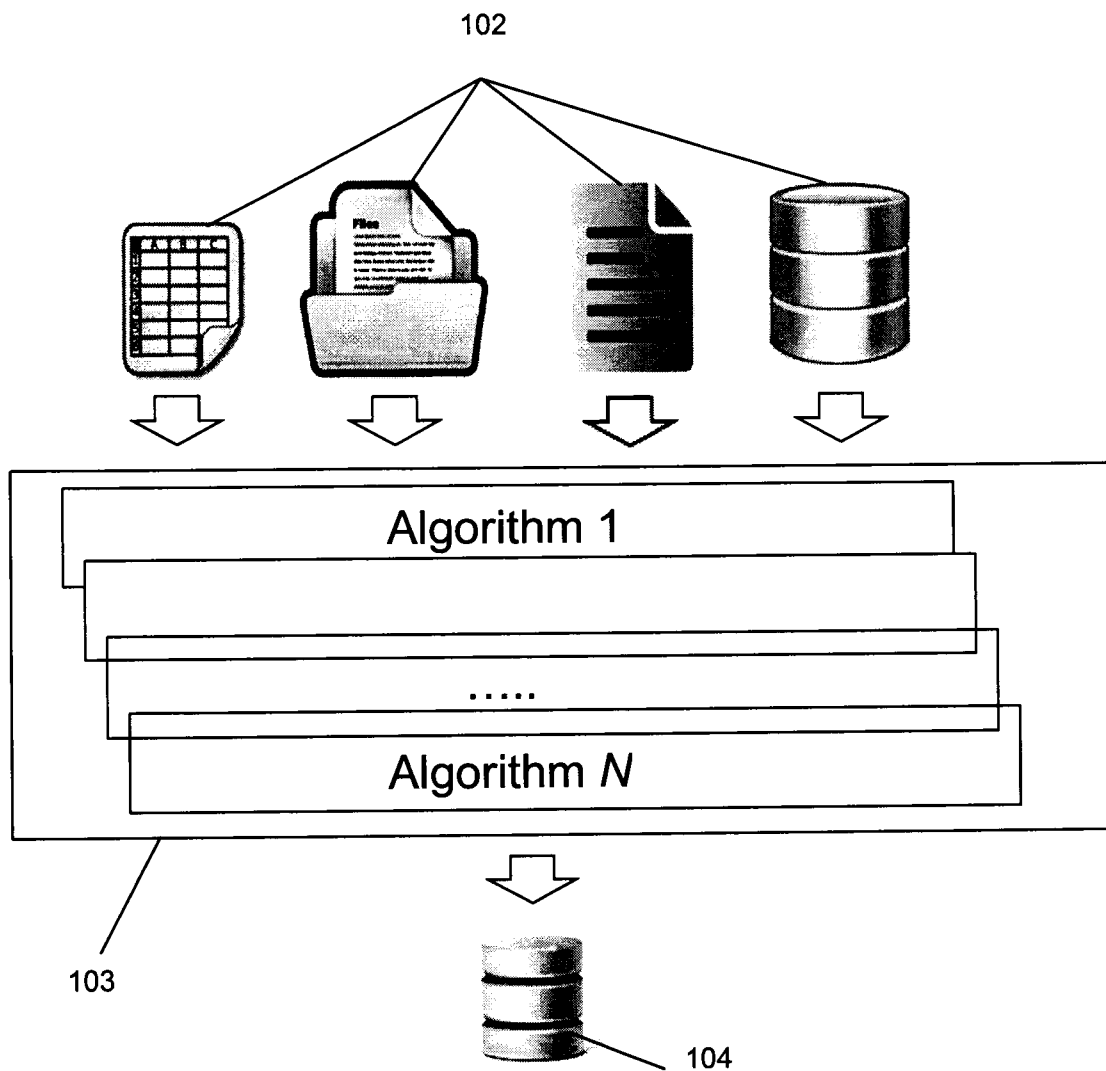


FIGURE 1

200

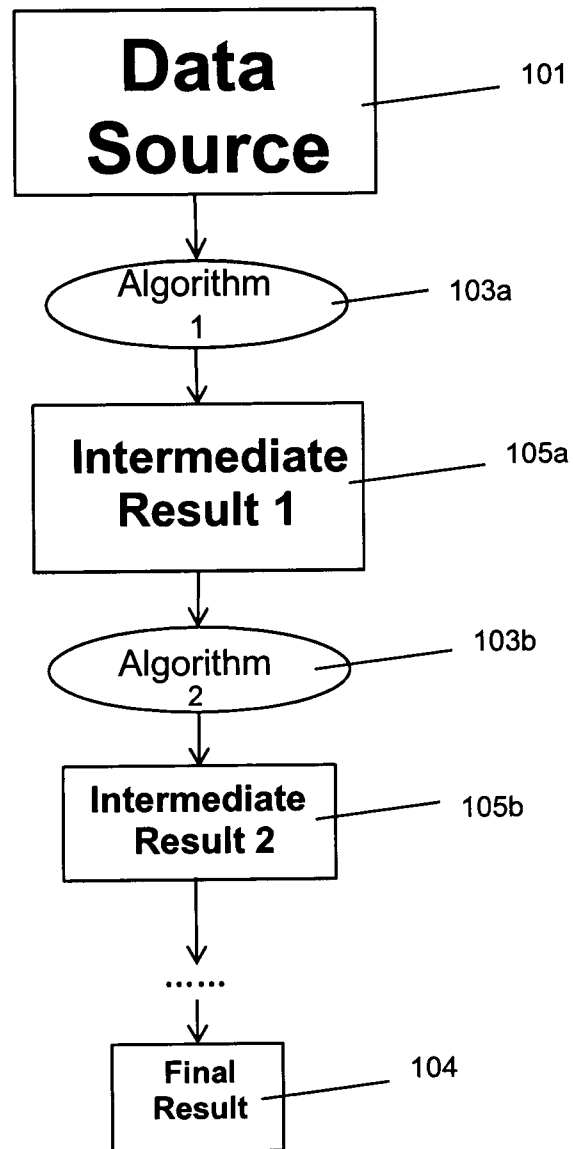


FIGURE 2

300

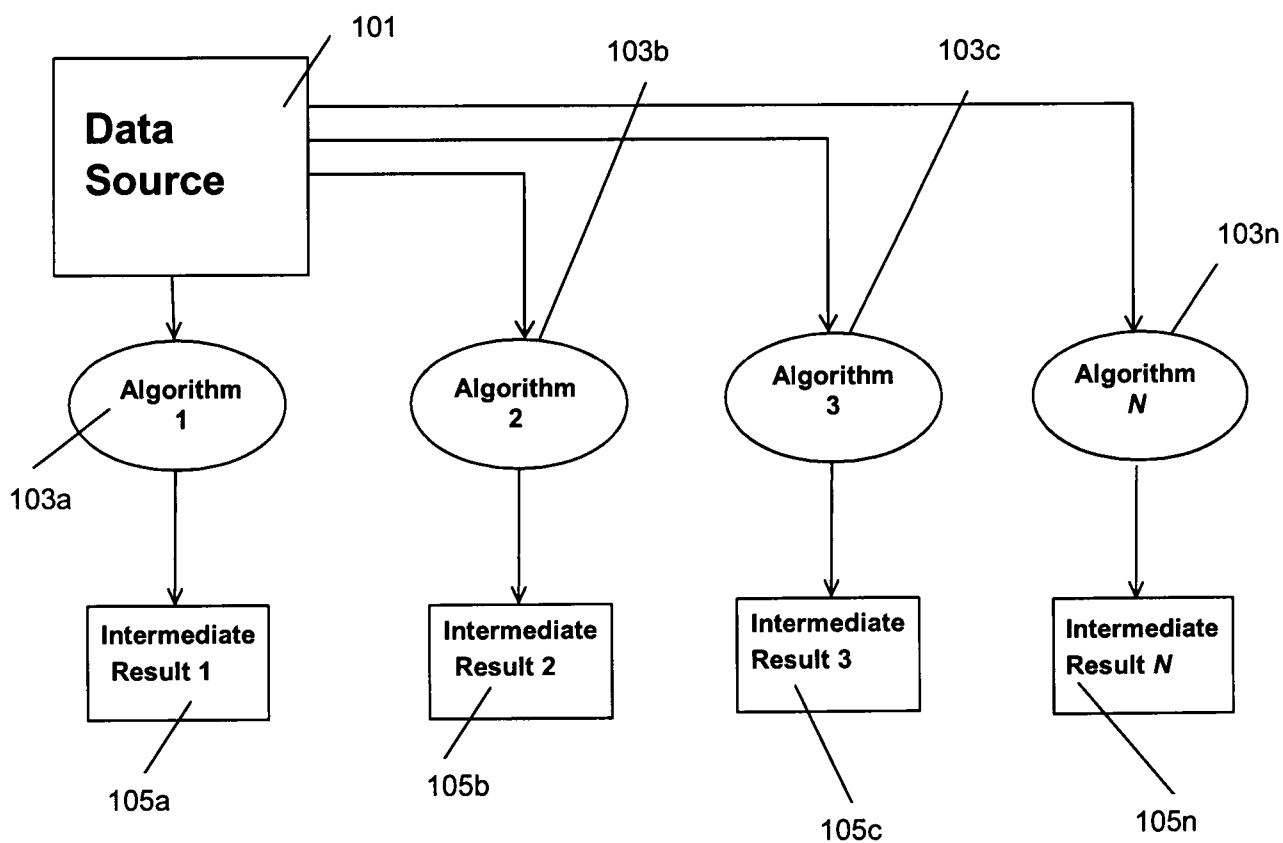


FIGURE 3

400

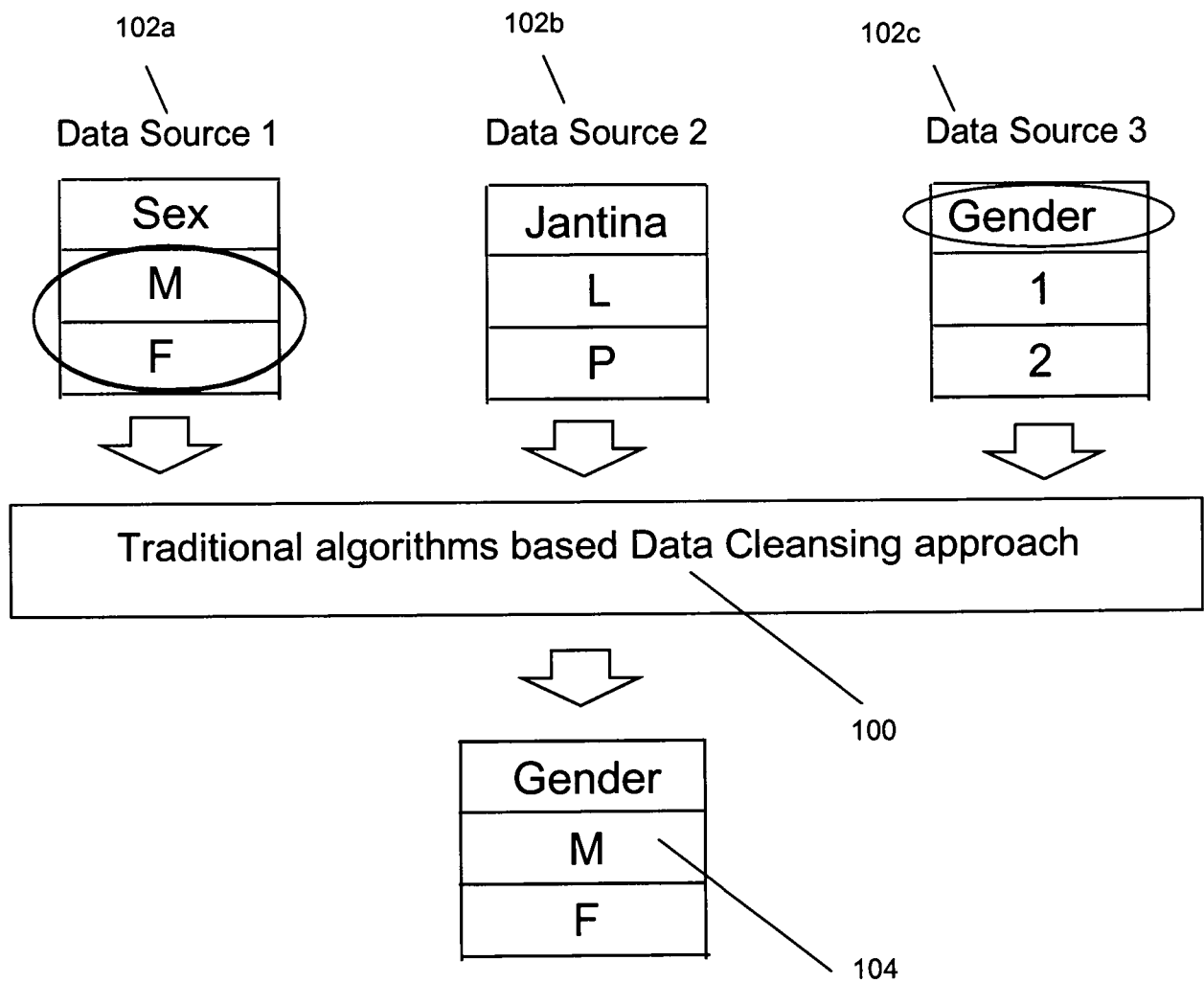


FIGURE 4

500

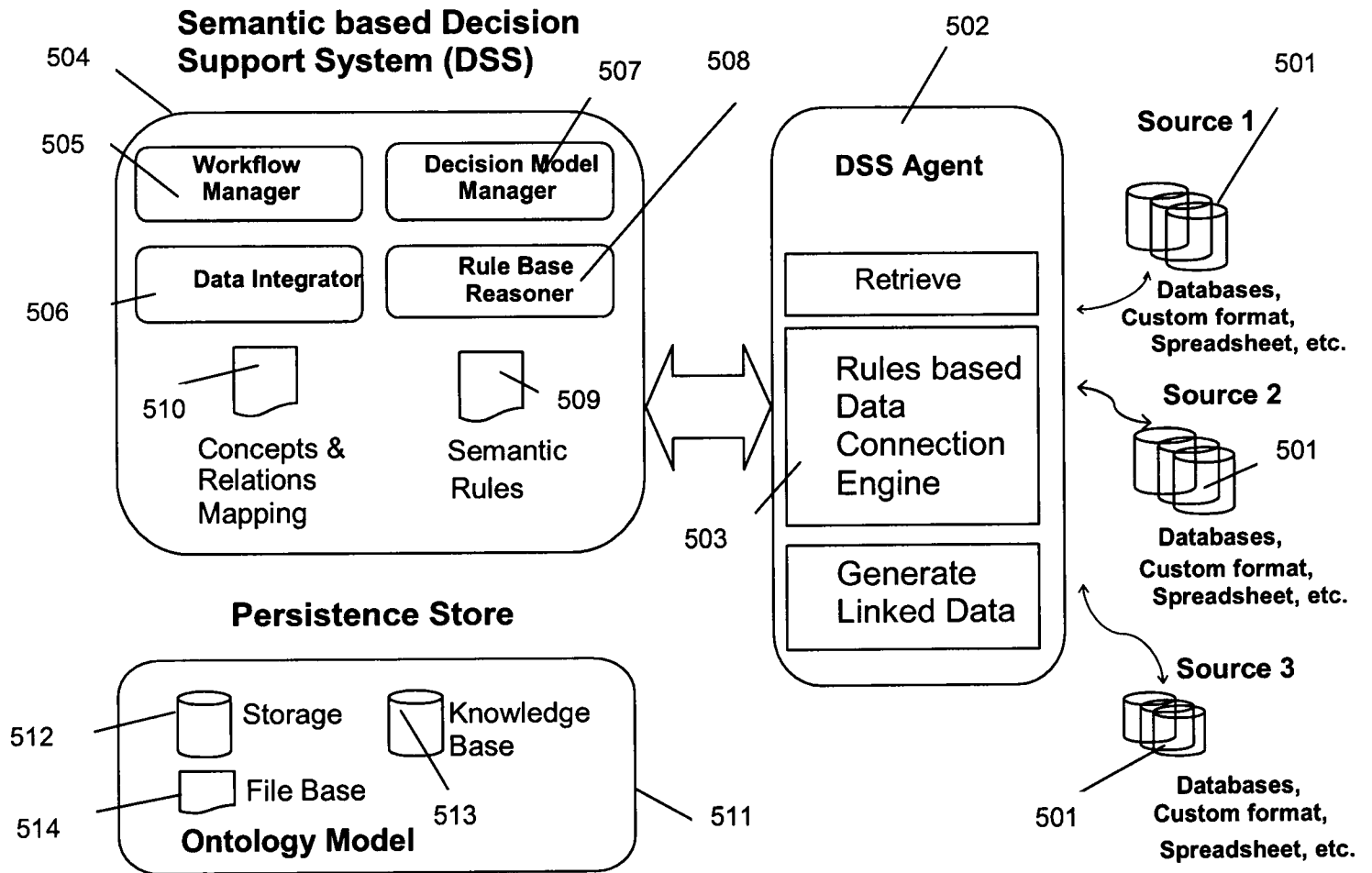


FIGURE 5

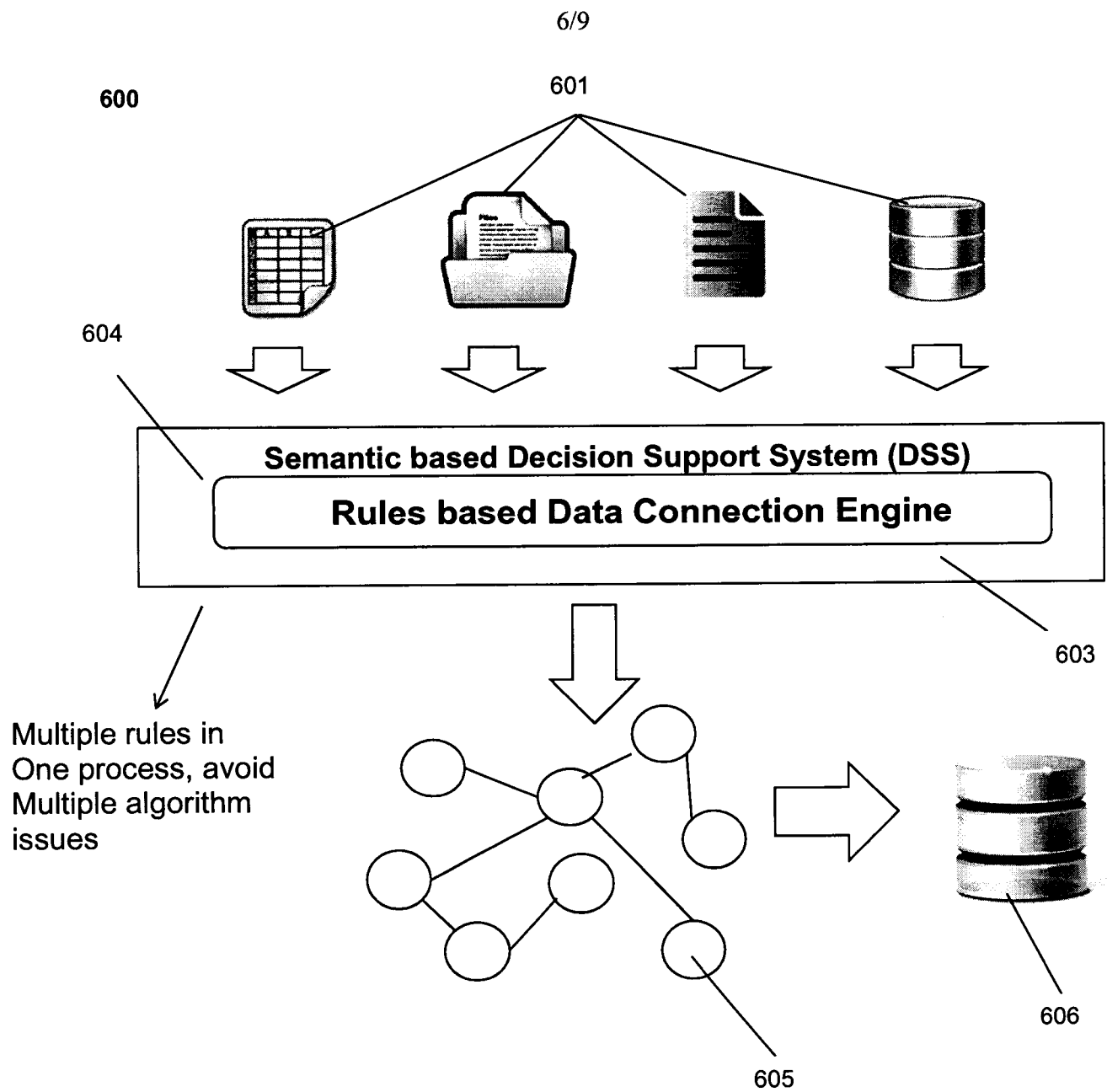


FIGURE 6

700

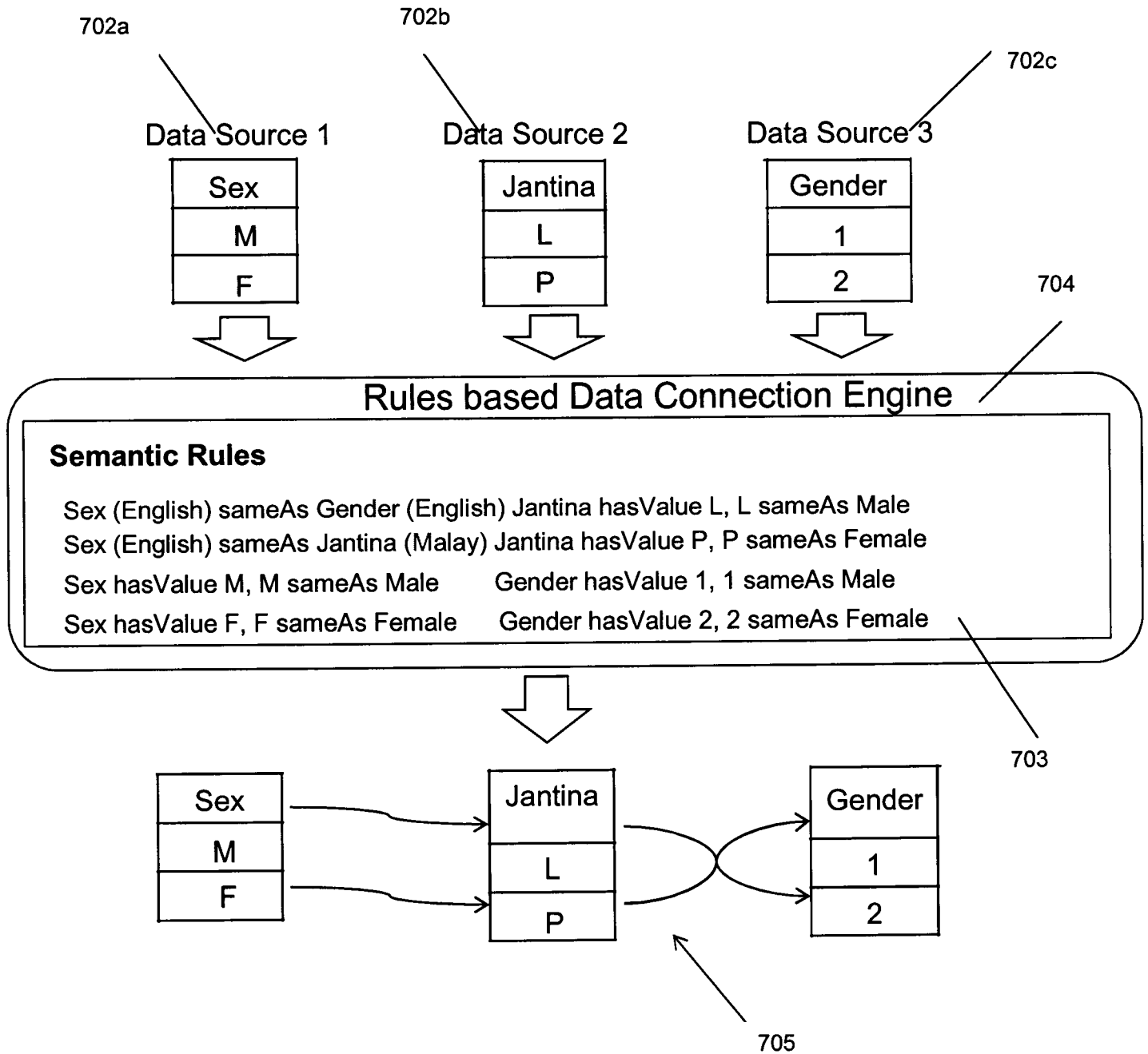


FIGURE 7

8/9

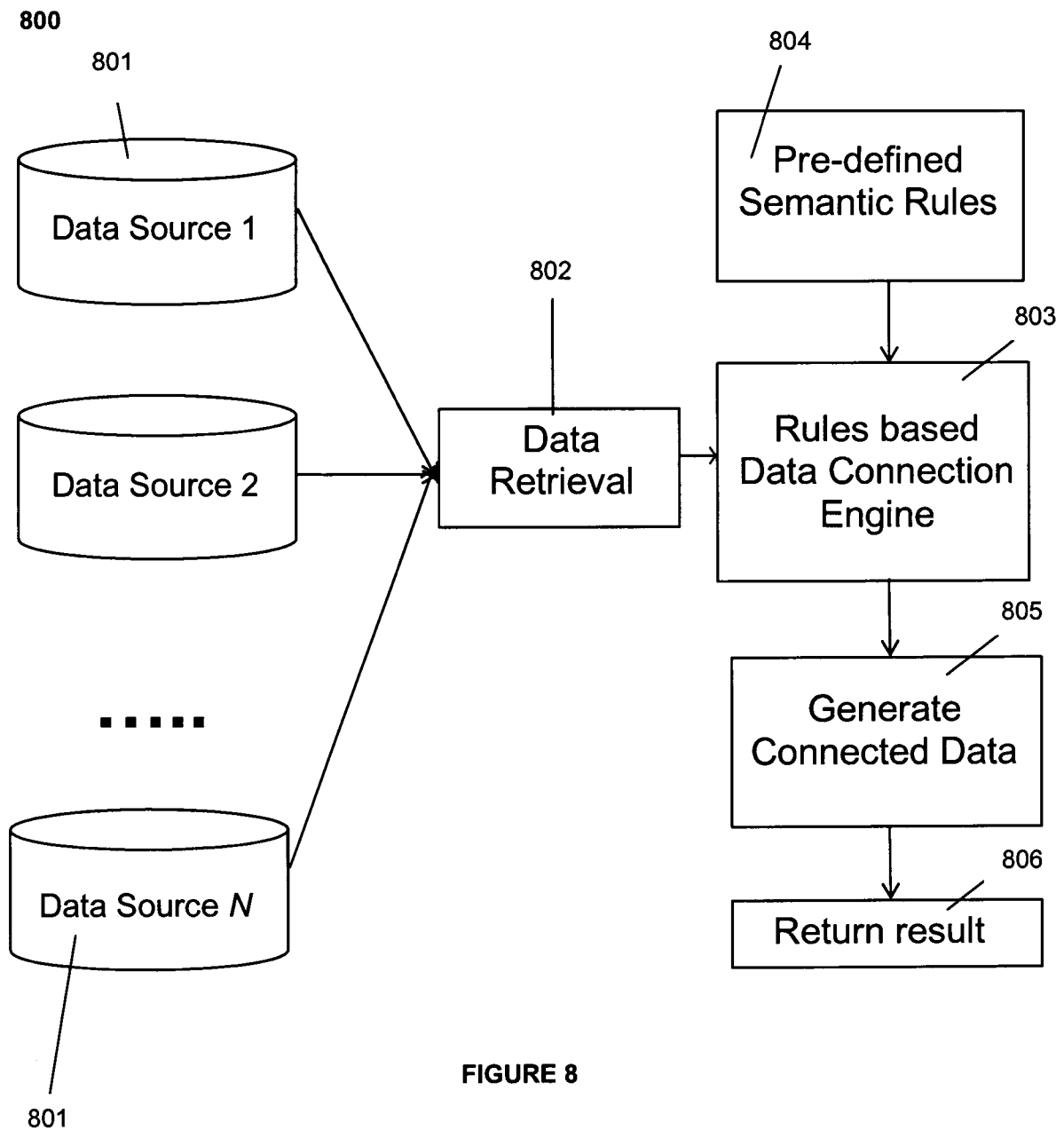


FIGURE 8

9/9

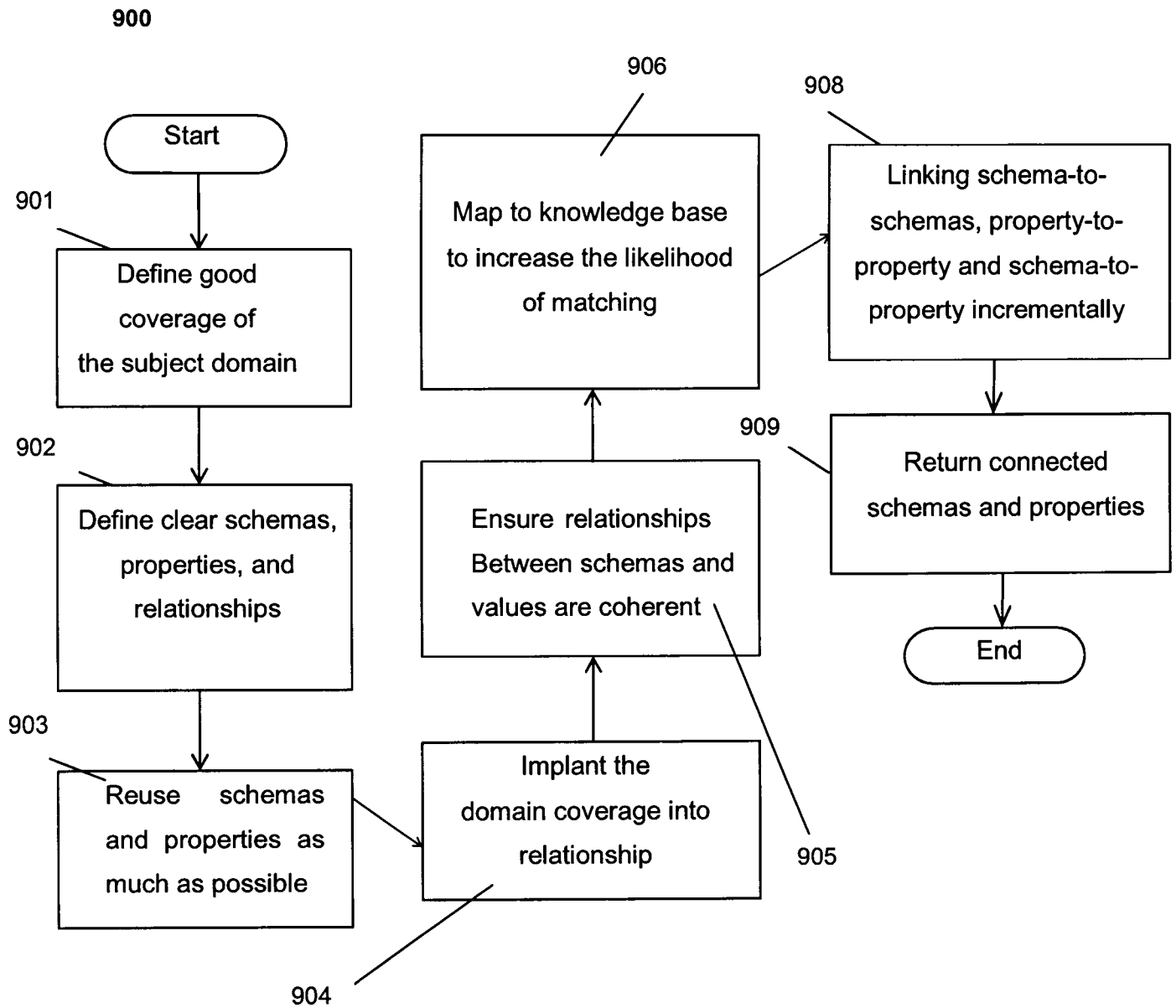


FIGURE 9

## INTERNATIONAL SEARCH REPORT

International application No  
PCT/MY2014/000149

A. CLASSIFICATION OF SUBJECT MATTER  
INV. G06N5/02  
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)  
G06N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages                                                                                                                                             | Relevant to claim No. |
|-----------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| X         | PHILIP A BERNSTEIN ET AL: "Generic Schema Matching, Ten Years Later", PROCEEDINGS OF THE VLDB ENDOWMENT, vol. 4, 3 September 2011 (2011-09-03), pages 695-701, XP55115656, the whole document                                  | 1-7                   |
| X         | MADHAVAN J ET AL: "GENERIC SCHEMA MATCHING WITH CUPID", PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON VERY LARGEDATA BASES, XX, XX, 31 January 2001 (2001-01-31), pages 1-10, XP001152140, the whole document<br>-----<br>-/- | 1-7                   |



Further documents are listed in the continuation of Box C.



See patent family annex.

\* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

26 March 2015

Date of mailing of the international search report

07/04/2015

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040,  
Fax: (+31-70) 340-3016

Authorized officer

Valencia, Erika

## INTERNATIONAL SEARCH REPORT

International application No

PCT/MY2014/000149

| C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT |                                                                                                                                                                                                                                                                                                                                                                                                                                          |                       |
|------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| Category*                                            | Citation of document, with indication, where appropriate, of the relevant passages                                                                                                                                                                                                                                                                                                                                                       | Relevant to claim No. |
| X                                                    | <p>CASTANO S ET AL: "A schema analysis and reconciliation tool environment for heterogeneous databases", DATABASE ENGINEERING AND APPLICATIONS, 1999. IDEAS '99. INTERNATIONAL SYMPOSIUM PROCEEDINGS MONTREAL, QUE., CANADA 2-4 AUG. 1999, LOS ALAMITOS, CA, USA, IEEE COMPUT. SOC, US, 2 August 1999 (1999-08-02), pages 53-62, XP010348659, DOI: 10.1109/IDEAS.1999.787251 ISBN: 978-0-7695-0265-6 the whole document</p> <p>-----</p> | 1-7                   |
| A                                                    | <p>JING LIN ET AL: "An Agent-Based Approach to Reconciling Data Heterogeneity in Cyber-Physical Systems", PARALLEL AND DISTRIBUTED PROCESSING WORKSHOPS AND PHD FORUM (IPDPSW), 2011 IEEE INTERNATIONAL SYMPOSIUM ON, IEEE, 16 May 2011 (2011-05-16), pages 93-103, XP031934701, DOI: 10.1109/IPDPS.2011.130 ISBN: 978-1-61284-425-1 the whole document</p> <p>-----</p>                                                                 | 1-7                   |