ICIS 1993 Proceedings

International Conference on Information Systems (ICIS)

1993

DIVERSITY WITH COOPERATION IN DATABASE SCHEMATA: SEMANTIC RELATIVISM

F. Saltor
Universitat Politenica de Catalunya

M. Garcia-Solaco
Universitat Politenica de Catalunya

Recommended Citation

Saltor, F. and Garcia-Solaco, M., "DIVERSITY WITH COOPERATION IN DATABASE SCHEMATA: SEMANTIC RELATIVISM" (1993). *ICIS 1993 Proceedings*. 53.
http://aisel.aisnet.org/icis1993/53

DIVERSITY WITH COOPERATION IN DATABASE SCHEMATA: SEMANTIC RELATIVISM

F. Saltor
M. García-Solaco
Universitat Politènica de Catalunya

ABSTRACT

Diversity is an asset, as long as it allows cooperation. In the case of information systems and databases, the data model used is a help or a hindrance for this cooperation of diverse views; this is characterized by the *semantic relativism* of the model.

We first analyze diversity within an information system, where cooperation is made possible by the use of external schemata; semantic relativism of the model of the database schema is the key factor. Then we discuss diversity between different information systems, where they cooperate through interoperability, by schema integration into federated schemata; semantic relativism of the canonical data model is shown to be determinant.

1. INTRODUCTION

Different people have diverse views of reality; they perceive and conceptualize the same portion of reality in diverse ways. We find not an absolute view, only relative views. Is this *diversity* good or bad? If it is an impediment to their cooperation, it is bad. The remedy is not uniformity, that is, imposing upon each of these people a single view, because then they would not feel at ease with this foreign view, and their work would be impaired.

If this diversity is not an impediment to their cooperation, and allows each of the people to work comfortably according to his/her view, it is an asset. Moreover, putting together all these views, one can get a richer conceptualization of the piece of reality.

The key word here is *cooperation*. Diversity with cooperation is the best approach. How can we get it? Diversity is a way of life, we do not need to pursue it. Cooperation, on the other hand, is not always easy to attain. People will cooperate if they are *able* to do it and *willing* to do it. The ability to cooperate depends on *technical* characteristics; within a computerized information system (IS), this ability is enhanced or impaired depending on the computer technologies used in the IS. The willingness to cooperate depends on *organizational* characteristics.

This paper discusses technical, not organizational, characteristics related to the cooperation of diversity. This does not mean that we consider these two kinds of characteristics as independent, or that we think that technical characteristics are more important than organizational characteristics; this only means that the scope of this paper is limited to technical aspects (for a discussion of some organizational aspects, see Goodhue, Wybo and Kirsch 1992). We therefore exclude considering different kinds of cooperation (from an organizational point of view) and taking into account security issues. Also, this paper analyzes technical characteristics at the level of issues and principles; technicalities are omitted here, and can be found elsewhere (Tsichritzis and Klug 1978; Saltor 1986, 1987; Sheth and Larson 1990; Saltor, Castellanos and Garcma-Solaco 1991; Castellanos, Saltor and Garcma-Solaco 1992).

The rest of the paper is organized as follows: Section 2 introduces *Semantic Relativism* as the feature corresponding to this ability to cooperate maintaining diversity. Diversity with cooperation within an IS is the subject of section 3, while diversity with cooperation between different information systems is discussed in section 4. We conclude in section 5.

2. DATA MODELS AND SEMANTIC RELATIVISM

2.1 Conception and Representation

We distinguish three *worlds*, or *realms*:

1) the real world, the *reality*;

2) the world of conceptions, or *conceptual world*; and

3) the world of representations, or *represented world.*

A *conception* is how a person conceives reality (or a part of reality); a *representation* is how a conception is represented by signs, the signs of a language (in the broad sense of the word).

The passage from the real world to the conceptual world is done by *perception* and *conceptualization*. The resulting conception is made of concepts and their relationships (abstractions). The issues involved are philosophical, psychological, sociological, etc., out of the scope of computer science.

The passage from the conceptual world to the represented world is done by *representation* (in the sense of the action of representing, while in the listing above "representation" was the result of representing). Representations are conventional, i.e., people involved agree by convention which signs to use and their relationships (*syntax*) and how structures of the language correspond to structures of the conceptual world (*semantics*).

The inverse passage from the represented to the conceptual world is done by *interpretation*. The result of interpreting a structure of the language is called its *meaning*, its significance.

There is never a direct passage between the real and the represented worlds: it is always done through the conceptual world. In this paper, we focus on the represented world, with references to the conceptual world. Real world per se is ignored (as are other worlds, such as the physical and social layers of Lindgreen 1990).

2.2 Representation Ability of Data Models

The previous explanation follows, in a very simplified way, general linguistic theory (see for instance Cherry 1966), but adopting the terminology that will be used throughout this paper. In the case of computer science, and according to ISO definitions, information belongs to the conceptual world, while *data* belongs to the represented world; i.e., information is *represented* by data, and the *meaning* of data is information.

A computerized IS is built upon a base made of data, adequately termed *database* (DB). A DB, that belongs to the represented world, is supposed to represent the conceptions of the participants of the IS, the users of the DB. Therefore, an important characteristic of a DB is its *representation ability*, i.e., how well can the DB represent those conceptions. The representation ability of a DB is given by its data model.

A *data model* is composed of structures, operations, and integrity constraints, or, following Brodie (1982), has static properties, dynamic properties and integrity rules.

The representation ability of a data model is composed of two factors, as presented by Saltor (1987): expressiveness and semantic relativism.

By *expressiveness* of a data model we mean the degree to which the model can directly represent (express in a natural way) any particular conception of the real world, no matter how complex this conception might be, and which concepts compose it. It approximately corresponds to what was called "semantic expressiveness" by Hammer and McLeod (1981), "conceptual naturalness" by Shipman (1981), and "modeling support" by Brodie (1982).

Expressiveness may be seen as composed of a structural part and a behavioral part. *Structural expressiveness* is the power of the structures of the model to represent concepts and to be interpreted as such concepts. *Behavioral expressiveness* reflects the power of the model to represent behaviors of concepts and to be interpreted correspondingly.

To illustrate the point, a model supporting generalization/specialization between superclasses and subclasses has more structural expressiveness than one not supporting it. A model that supports aggregation/decomposition between complex data objects and their constituent data objects has more structural expressiveness than one that does not.

For example, the relational model has no direct support for generalization or aggregation, i.e., has no constructs that can be directly interpreted as such abstractions, and therefore has less structural expressiveness than some extended Entity Relationship or Object Oriented models that have such constructs.

A model that supports not only generic operations of its structures and integrity constraints inherent to the model but also the definition of new operations and integrity constraints has more behavioral expressiveness than a model not supporting this feature.

The *semantic relativism* of a data model is the degree to which the model can help in representing not only one but many *diverse* conceptions of the same real world, and at the same time allow them to cooperate. This concept is the central topic of this paper. We will look at examples later on. It is similar to "semantic relativism" of Brodie (1982) and Spaccapietra and Parent (1990) and to "relativism" (Hammer and McLeod 1981).

To support diversity with cooperation, a system needs not only a data model with a high degree of semantic relativism, but also an adequate architecture, as will be shown in the next two sections. The data model, however, is very important, as we will see.

3. DIVERSITY WITH COOPERATION WITHIN AN IS: EXTERNAL SCHEMATA

Different subunits and employees of an organization perceive and conceptualize the organization and its environment in diverse ways, according to their relative points of view. Considering, in this section, a single IS for the whole organization, each of these conceptions (in the conceptual world) is represented by an schema (in the represented world). For these users to cooperate, there shall not be a separate DB for each of these schemata, with data duplication and other redundancies; there should be a common DB, storing each data element just once, according to a more general schema that encompasses all those relative schemata and avoiding redundancies.

This general schema is called "conceptual schema" by ANSI/SPARC (Tsichritzis and Klug 1978) (even if it belongs to the represented, not the conceptual world), "database schema" in the ISO terminology of (van Griethuysen 1982), and is the representation of the "information model" of (Lindgreen 1990); in this paper, we will use the term *database schema*. The relative schemata of the users are called *external schemata* by ANSI/SPARC (Tsichritzis and Klug 1978). When we speak of the data model of a DB, we mean the model of its database schema.

For the common DB to support all users' conceptions, i.e., to implement semantic relativism, its DB Management System (DBMS) must have an architecture that supports all these external schemata, and mechanisms for their *derivation* from the database schema. Some users may need not only to query but also to update the DB through their external schemata, a research topic called *view updating*; the mechanisms of current DBMSs provide little support in this respect.

To construct the DB, a DBMS is selected, and the database schema is designed. This design may be done by the *integration* of the external schemata into the database schema (note that this process is the opposite of the derivation process found before). Once the DBMS has been selected, the database schema has been designed, and the DB has been populated with data, this DB is imposed on all users. If the semantic relativism of the DBMS is not good enough, it may not be possible to design a database schema, and to define a derivation from it, to obtain precisely the external schema representing a given user's conception.

The power to derive external schemata from the database schema is therefore the measure of semantic relativism of a DBMS and of its data model (Saltor 1986).

The semantic relativism of pre-relational models is quite limited, because they only allow external schemata that are strict subsets of the database schema. Relational DBMSs may use the whole power of relational calculus or algebra to derive external schemata ("views"), and have therefore a high degree of semantic relativism (Saltor 1986): not only do they allow views with "virtual attributes" (computed from several stored attributes, such as Amount = Quantity x Price) and "aggregates by reduction" of an attribute (totals, averages, counts, maximum and minimum values), but they also support views aggregating two or more entities, hiding if necessary how they are related (join and project), etc.

The Entity Relationship (E-R) model, even when equipped with operations (such as the Burgundy algebra of Parent and Spaccapietra 1985), was shown in Saltor (1986) to have less semantic relativism that the relational model. This is also generally true for extensions of the E-R model (Elmasri, Weeldreyer and Hevner 1985; Teorey, Yang and Fry 1986; Parent and Spaccapietra 1989).

Support of views in Object Oriented (OO) Databases is a promising research topic (Abiteboul and Bonner 1991; Scholl and Schek 1991) not commonly available in existing OO-DBMSs.

4. DIVERSITY WITH COOPERATION BETWEEN DIFFERENT INFORMATION SYSTEMS: INTEROPERABILITY

The cooperation between several ISs may arise in a number of cases. Different organizations, each with its own IS, may want to cooperate (subsidiaries of a common parent company, states of a federal country, government agencies, countries forming a common market, etc.). An organization might have several ISs, for example one per division, developed independently. Two companies may merge, or a takeover could take place, and keeping their respective ISs and having them cooperate may be preferable to their substitution by a new, common IS.

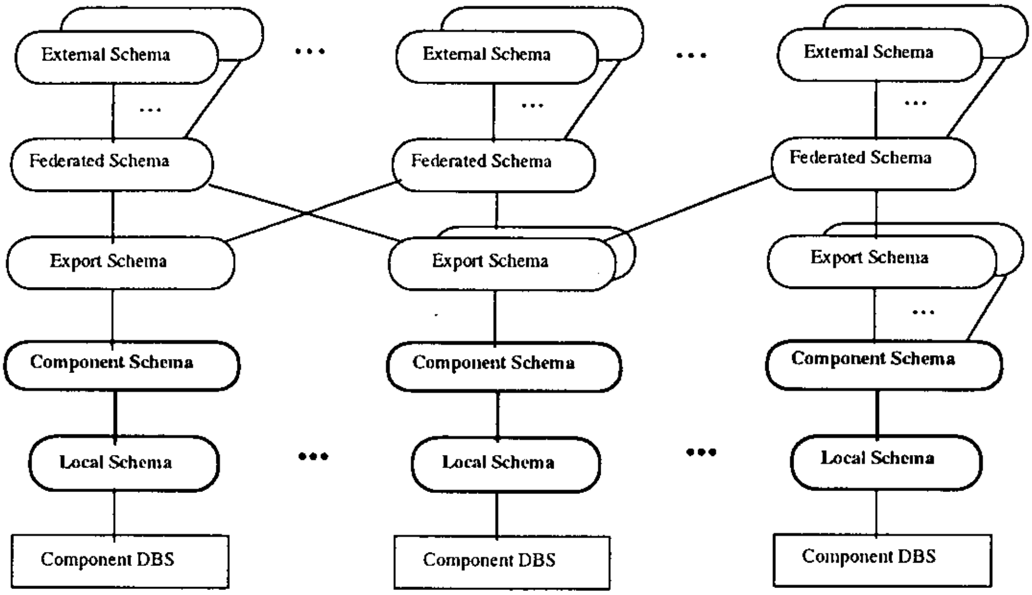

Figure 1. Five Level Schema Architecture of a FDBS

In each of these cases, we find different ISs, designed independently and operating autonomously, that are bound to cooperate. We assume that the technical form of cooperation is *interoperability*, supporting *integrated access* to the collection of ISs. This means that a user is able to ask a single query (one access), and receive a single, consolidated answer; the individual ISs have cooperated to produce this answer. Which of the ISs have provided which data may be hidden to the user and he has the feeling of accessing a single IS. Alternatively, data may be tagged with an identification of the IS supplying them (called "source tagging" in Wang and Madnick 1990), depending on the needs of the user.

Before analyzing the role of semantic relativism of data models in this context, let us present the terminology and the architecture we will be using.

4.1 Terminology and Architecture

We will say that the individual ISs, called *component ISs*, form a federation, or a *federated IS*. Considering their respective DB systems, a *federated DB system* is obtained through the interoperability of the *component DB systems*. The federated DB system (FDBS) has no data of its own: it answers queries by accessing the component DB systems. It is a layer of software placed on top of the DBMSs of the component DBs (other software and hardware needed are out of the scope of this paper).

To support integrated access, the FDBS must have an adequate architecture and a convenient data model. To discuss how semantic relativism achieves diversity with cooperation, we must center on a given architecture. We will follow the five level schema architecture and the terminology of Sheth and Larson (1990) shown in Figure 1.

First, a *canonical* data model (CDM), common to the whole federation, must be adopted. The database schemata of the component DBs (*local schemata*) are transformed from their native models to the CDM, giving *component schemata*. Each component schema is filtered into one or more *export schemata*. From export schemata of different component DBs, a *federated schema* is constructed. This process is called *schema integration*; several federated schemata can exist in a federation. Finally, from a federated schema a number of *external schemata* are derived, for different users (or categories of users) of the federated IS.

A query using an external schema will be mapped to its federated schema, decomposed into subqueries to the component ISs concerned, translated to their local schemata, and submitted to the corresponding DBMSs. These provide (sub)results, which are translated, consolidated, formatted, and presented to the user. This process is not covered in this paper.

When a federation is formed, or when an IS is to enter an existing federated IS, a *negotiation* process takes place. Each IS negotiates which of its own data it makes accessible (it "exports") to the federation (to which categories of users of the federation), and which part of these data, if any, it allows not only to be read but also to be updated (by which categories of users). It is clear, therefore, that an IS may prevent part of its data from access by other ISs this is the role of the export schemata. Negotiation may include to which point, if any, the autonomy of the IS is substituted by an interdependence with the other ISs, without compromising its support for local, preexisting users. We do not discuss negotiations in this paper.

Since each component IS was designed independently of each other, the diversity of decisions will have led to a number of heterogeneities. Examples of *systems heterogeneities* are different CPUs, operating systems, data models and languages, DBMSs, communication protocols. *Data heterogeneities* are due to the different conceptions by the designers of the component ISs (*semantic* heterogeneities) and to the different representations of these conceptions in the respective data models and DBMSs (*syntactical* heterogeneities), and are the subject of this section.

The use of a common CDM, and the transformation from local schemata into component schemata, solve the problem of syntactic heterogeneities. Semantic heterogeneities are dealt with in the schema integration process, by which export schemata are integrated into federated schemata.

We have seen that this five level architecture supports integrated access to the federated IS. Three layers in this architecture, namely component schemata, export schemata and federated schemata, are expressed in the CDM (external schemata may be expressed in different models, as we will see). How well this architecture solves data heterogeneities depends on the characteristics of the CDM, more particularly on the two factors of its representation ability seen in section 2.2: expressiveness and semantic relativism.

4.2 Expressiveness of the CDM

A CDM must have an expressiveness equal or greater than any of the native models of the component DBs that are going to interoperate, in order to capture the semantics already explicitly expressed in their local schemata. Moreover, it should support additional semantics, implicit in the data values and not expressed in the local schemata due to the poor expressiveness of their native models. These additional semantics may be discovered and made explicit through a semantic enrichment process, as in Castellanos (1993).

Technical characteristics of the expressiveness of a data model that make it suitable as the CDM of a federated IS are detailed in Saltor, Castellanos and Garcma-Solaco (1991). In particular, semantic models, in the sense of Peckham and Maryanski (1988), and object oriented models, satisfying Atkinson et al. (1990), are more expressive than relational and extended E-R models.

4.3 Semantic Relativism of the CDM

In order to support diversity with cooperation in the context of interoperability, two processes are crucial: the integration of export schemata into federated schemata and the derivation from these of external schemata. Schema integration makes possible the cooperation between different ISs with diverse database schemata, while the second process supports cooperation between users of the federated IS (probably previous users of one of the component ISs) with very diverse external schemata.

Schema integration is the most difficult process in the formation of a federated IS. It must overcome all kinds of semantic heterogeneity: naming conflicts, structural and schematic conflicts, unit conflicts, domain conflicts (including scale and precision), etc. Some of these problems are not solved at this stage, and a lot of research is going on in this area, as pointed out by Bukhres, Elmagarmid and Mullen (1992).

The process of schema integration becomes less difficult if the CDM is very expressive, and export schemata make use of this expressiveness and are semantically rich, and if the CDM has operations to support their integration. These integration operations contribute to the semantic relativism of the CDM.

Another factor that may appear when integrating export schemata is that different users may have conceptions that are not subsets of a more general conception, but that diverge in some respect. An example (Sheth and Larson 1990) is the integration of colors of shoes from two DBs, DB1 and DB2: ConceptionA sees as "cream" what is cream in DB1 and what is tan in DB2, while ConceptionB considers "cream" what is tan or cream in DB1 and what is tan or white in DB2.

This is called *multiple semantics* by Sheth and Larson and must be supported by the Federated IS. One way to handle it is by having one federated schema for each semantic conception (Sheth and Larson 1990), i.e., a federated schema for ConceptionA and another for ConceptionB. An alternative is to allow a single federated schema to support multiple semantics (for example, through the discriminants of Garcma-Solaco and Saltor 1991) and differentiate the respective conceptions at the external schema level, i.e., an external schema for ConceptionA and another for ConceptionB. One of the advantages of the second architecture lies in having a smaller number of federated schemata; on the other hand, it requires a CDM supporting multiple semantics at the federated schema level, both in structures and in operations, able to derive external schemata for each semantic conception.

Derivation of external schemata needs a greater flexibility in this context than in the case of a single IS. Two problems related to the architecture of the FDBS and not well solved are the *view updating* problem, more difficult here than in the single IS case seen in section 3, and the *transaction management* problem, particularly for updating transactions. They are beyond the scope of this paper.

Considering factors related to the canonical model, not only do the operations of the CDM have to support derivation of any external schema desired, the user may want his external schema expressed in a given data model, different from the CDM, because it is the model he is familiar with — it may be the model of the component IS he used to work with. Multimodel support at the external schema level (a research topic not well solved) is therefore a factor in the architecture and in the semantic relativism of the CDM.

In summary, characteristics of the semantic relativism of a data model that make it more suitable as the CDM of a federated IS include integration operations, support of multiple semantics, and power to derive external schemata with multimodel support. These characteristics are technically detailed in Saltor, Castellanos and Garcma-Solaco. In that paper, an analysis of data models as possible candidates for the CDM of a federated IS, considering their semantic relativism, is performed. It concludes that relational and extended E-R models fail to satisfy some of the characteristics. Functional models and TAXIS (Mylopoulos, Bernstein and Wong 1980) but not other semantic models are better placed. OO models satisfying Atkinson et al. (1990), if equipped with good operations and supporting views, such as those in Abiteboul and Bonner (1991) and Scholl and Schek (1991), are found to be the best in this respect.

5. CONCLUSIONS

Diversity is not the contrary of cooperation, but of uniformity. Cooperation is not incompatible with diversity, but with separation. If it is possible to maintain diversity, without imposing uniformity, and at the same time to allow cooperation, then we will have the best approach.

In the case of information systems and databases, the characteristics of data models are very important. We have presented *semantic relativism* as a key factor of a data model and of its representation ability.

In the case of a single information system, the cooperation between different users with diverse conceptions of the organization and its environment is made possible through the external schema mechanism of the DBMS. The data model of the DBMS, i.e., the model of the database schema, makes this cooperation easier or more difficult, according to its semantic relativism. We have shown that relational models, as well as Object Oriented models supporting views, have a high degree of semantic relativism, and therefore the corresponding systems have a good support for this cooperation.

In the context of the cooperation between different information systems, where diversity is much more extended and pervasive, the architecture and the canonical data model are key factors. Technical problems not yet solved include overcoming semantic heterogeneities in the schema integration process, as well as view updating and transaction management for updating users.

Concerning the canonical data model, we have presented characteristics that it should have to contribute to its semantic relativism, and therefore to the cooperation between the information systems and their users. Object Oriented models, if equipped with a good support for views, are best placed in this respect, and are therefore good candidates to support diversity with cooperation in federated information systems.

Therefore, diversity with cooperation depends to a large extent, in both cases, on the semantic relativism of the data model used.

6. ACKNOWLEDGMENTS

This research has been partially supported by PRONTIC project TIC89/0303. We thank M. Castellanos and the anonymous reviewers for their comments to a previous version of this paper.

7. REFERENCES

Abiteboul, S., and Bonner, A. "Objects and Views." *Proceedings of the ACM SIGMOD Conference.* ACM SIGMOD Record, Volume 20, Number 2, 1991, pp 238-247.

Atkinson, M.; Bancilhon, F.; DeWitt, D.; Dittrich, K.; Maier, D.; and Zdonik, S. "The Object Oriented Database Systems Manifesto." *Proceedings of the International Conference on Deductive and Object Oriented Databases*, Amsterdam: Elsevier, 1990.

Brodie, M. L. "On the Development of Data Models." In M. L. Brodie, J. Mylopoulos and J. W. Schmidt (Eds.), *On Conceptual Modelling*, Berlin: Springer Verlag, 1982.

Bukhres, O.; Elmagarmid, A.; and Mullen, J. "Object-Oriented Multidatabases: Systems and Research Overview." *Proceedings of the International Conference on Information and Knowledge Management*, Baltimore, 1992.

Castellanos, M. "A Methodology for Semantically Enriching Interoperable Databases." *Proceedings of the Eleventh British National Conference on Databases*, Keele, 1993.

Castellanos, M.; Saltor, F.; and Garcma-Solaco, M. "A Canonical Model for the Interoperability Among Object-Oriented and Relational Databases." In Ozsu, Dayal and Valduriez (Eds.), *Distributed Object Management* (Proceedings of the International Workshop on Distributed Object Management, Edmonton, Canada, 1992), Morgan Kaufmann, 1993.

Cherry, C. *On Human Communication*. Cambridge, Massachusetts: MIT Press, 1966.

Elmasri, R.; Weeldreyer, M.; and Hevner, A. "The Category Concept: An Extension to the Entity Relationship Model." *Data and Knowledge Engineering*, Volume 1, Number 1, 1985, pp. 75-116.

Garcma-Solaco, M., and Saltor, F. "Discriminated Operations for Interoperable Databases." *Proceedings of the First International Workshop on Interoperability in Multidatabase Systems*, Kyoto, 1991.

Goodhue, D.; Wybo, M.; and Kirsch, L. "The Impact of Data Integration on the Costs and Benefits of Information Systems." *MIS Quarterly*, Volume 16, Number 3, September 1992, pp. 293-311.

Hammer, M., and McLeod, D. "Database Description with SDM: A Semantic Database Model." *ACM Transactions on Database Systems*, Volume 6, Number 3, 1981, pp. 351-386.

Lindgreen, P. (Ed). *A Framework of Information Systems Concepts* (Interim Report). University of Nijmegen, 1990.

Mylopoulos, J.; Bernstein, P.; and Wong, H. "A Language Facility for Designing Database Intensive Applications." *ACM Transactions on Database Systems*, Volume 5, Number 2, 1980, pp. 185-207.

Peckham, J., and Maryanski, F. "Semantic Data Models." *ACM Computing Surveys*, Volume 20, Number 3, 1988, pp. 153-189.

Parent, C., and Spaccapietra, S. "About Entities, Complex Objects and Object Oriented Data Models." In Falkenberg and Lindgreen (Eds.), *Information Systems Concepts: An In-depth Analysis*. Amsterdam: North Holland, 1989.

Parent, C., and Spaccapietra, S. "An Algebra for a General Entity-Relationship Model." *IEEE Transactions on Software Engineering*, Volume 11, Number 7, 1985, pp. 634-643.

Saltor, F. "On the Power to Derive External Schemata from the Database Schema." *Second International Conference on Data Engineering*, Los Angeles, 1986.

Saltor, F. "Semantic Relativism in Databases: The Case of RM/T." Facultat d'Informatica RR 18/87, UPC, Barcelona, 1987.

Saltor, F.; Castellanos, M.; and Garcma-Solaco, M. "Suitability of Data Models as Canonical Models for Federated Databases." *ACM SIGMOD Record*, Volume 20, Number 4, December 1991, pp. 44-48 (Special Issue on Semantic Issues in Multidatabase Systems).

Sheth, A., and Larson, J. "Federated Database Systems for Managing Distributed, Heterogeneous, and Autonomous Databases." *ACM Computing Surveys*, Volume 22, Number 3, 1990, pp. 183-236.

Shipman, D. "The Functional Data Model and the Data Language DAPLE." *ACM Transactions on Database Systems*, Volume 6, Number 1, 1981, pp. 140-173.

Spaccapietra, S., and Parent, C. "View integration: A Step Forward in Solving Structural Conflicts." DI, EPF Lausanne, 1990.

Scholl, M., and Schek, H. J. "Supporting Views in Object Oriented Databases." *IEEE-CS Bulletin on Data Engineering*, Volume 14, Number 2, 1991, pp. 43-47.

Tsichritzis, D., and Klug, A. "The ANSI/X3/SPARC DBMS Framework." *Information Systems*, Volume 3, Number 4, 1978.

Teorey, T. J.; Yang, W.; and Fry, J. P. "A Logical Design Methodology for Relational Databases Using the Extended Entity Relationship Model." *ACM Computing Surveys*, Volume 18, Number 2, 1986, pp. 197-222.

van Griethuysen, J. (Ed.). *Concepts and Terminology for the Conceptual Schema and the Information Base*. ISO/TC97/SC5/WG3, 1982.

Wang, Y., and Madnick, S. "A Polygen Model for Heterogeneous DBS: The Source Tagging Perspective." *Proceedings of the Sixteenth Very Large Data Base Conference*, Brisbane. Morgan Kaufmann, 1990.