Collaborative Information Management System for Science Domains

Ozgul Unal
University of Amsterdam, The Netherlands

Ersin C. Kaletas
University of Amsterdam, The Netherlands

Hamideh Afsarmanesh
University of Amsterdam, The Netherlands

H. Hakan Yakali
University of Amsterdam, The Netherlands

Louis O. Hertzberger
University of Amsterdam, The Netherlands

NEED FOR COLLABORATION

With the increasing need for collaboration in different science domains, a lot of research activities are now focused on the mechanisms and infrastructures supporting advanced collaborations among pre-existing, distributed, heterogeneous, and autonomous organizations. Collaborating organizations typically share some common objectives and in order to achieve them they need to share their information and resources. One prominent requirement is to access each other's data or databases through a secure infrastructure. Biodiversity is one such science domain. Challenges in biodiversity information management are being addressed in the project ENBI (European Network for Biodiversity Information) (ENBI, n.d.). A summary of the information management challenges in different science domains is given in the second section, Information Management Challenges in Science Domains.

In this context, the CIMS introduces promising solutions to cope with these challenges. In general terms, CIMS refers to the set of components and mechanisms that together constitute a generic information manipulation framework to support the interoperation and data sharing among collaborating members (Guevara-Masis, Unal, Kaletas, Afsarmanesh, & Hertzberger, 2004).

The proposed CIMS involves three main paradigms and technologies, consisting of: (1) federated database architecture, (2) virtual organizations paradigm, and (3) grid technology. Introductory information about these technologies and paradigms and their major benefits are covered in the third section, paradigms and technologies.

INFORMATION MANAGEMENT CHALLENGES IN SCIENCE DOMAINS

Below, a summary of challenges and requirements of information management in science domains is provided. The points addressed here correspond to the main motivations behind the need for a CIMS.

- **Distributed and Heterogeneous Databases:** One of the challenges in many science domains is to have transparent access to distributed and heterogeneous data sources. Existing organizations typically employ different data structures depending on their specific needs. A network of organizations must consider such differences for providing effective mechanisms to integrate or inter-link and homogeneously access such databases.
- **Autonomous Organizations:** Organizations from a variety of science domains represent autonomous nodes running independently of each other. Each organization must be able to autonomously decide to share a part of their local resources or services with certain other specific organizations, based on some agreements.
- **Need for Collaboration:** It has become more clear that the collaboration between different organizations, activities, and users in science domains is important for an improved understanding and thus for achieving better results in the domain. However, most organizations do not want to actively cooper-

ate, because of for example the sensitiveness of some data that they have. Therefore, mechanisms supporting collaboration among organizations and at the same time taking these kinds of data into account are needed. With the existence of such a coordinating mechanism, organizations can more easily decide to collaborate.

- **Security and Access Rights:** The issues of trust, inter-organizational agreements, etc. play an important role in resource sharing among organizations from different science domains. Access rights and visibility levels must be taken carefully into account especially considering sensitive data. A hierarchy of visibility levels and access rights needs to be defined for different users or user roles.
- **Performance Requirements:** The amount of online data from different science domains, such as biology, physics, and astronomy, is increasing at a high pace. Considering the need for sharing these data with others and making them accessible through a network, the demands for performance and robustness are high. Furthermore, some modeling and analysis activities require computationally intensive algorithms. Thus, an infrastructure providing high performance distributed resource and data management facilities is needed.

PARADIGMS AND TECHNOLOGIES

In order to address the information management challenges identified in the second section, information management challenges in science domains, we propose to use a CIMS, which is a combination of federated database architecture, virtual organizations paradigm, and grid services.

Federated Database Architecture

Federated database architecture is a kind of distributed information management system. It provides necessary mechanisms for access to distributed, heterogeneous, and autonomous databases and enables sharing and exchange of information among collaborating partners (Afsarmanesh, Garita, Ugur, Frenkel, & Hertzberger, 1999; Busse, Kutsche, Leser, & Weber, 1999; Garita, Ugur, Frenkel, Afsarmanesh, & Hertzberger, 2000; Sheth & Larson, 1990). Federated database architecture facilitates the definition of integrated database schemas taking the visibility issues into account. Every node in the federation maintains autonomy on the local data, uses its own data model, and defines on it a set of export schemas for each other node based on the access rights. These export schemas are then imported into the integrated schema definitions. The federated architecture enables a flexible definition of access rights. Furthermore, by means of federated query processing user can transparently access to distributed and heterogeneous data using a single interface, which hides the physical data location.

Virtual Organizations

The virtual organization paradigm has drawn considerable attention in recent years, as the large and small organizations are more and more faced with the need to collaborate and to combine their capabilities, resources, and expertise, in order to remain competitive in a highly aggressive global market, as well as to be able to provide value-added services and products. Given the complexity and multidimensionality nature of the virtual organizations, there are still no standard definitions for virtual organizations. In addition to the lack of a common definition for virtual organization, a number of different terms are competing in the literature that refer to similar concepts or to its associated branches. For instance, examples of related terms to virtual organizations include: Virtual enterprise (VE), extended enterprise, supply chain management, enterprise clusters, and networked enterprises among others (Garita, 2001), where the closest concept to the current definition of virtual organizations is the one of VEs. In 1999, the definition of a VE is given by Camarinha-Matos and Afsarmanesh (1999) as follows:

A Virtual Enterprise (VE) is a temporary alliance of enterprises that come together to share skills, core competencies and resources in order to better respond to business opportunities, and whose cooperation is supported by computer networks. (p. 4)

Another definition for virtual organizations is provided by European Commission (2002):

The virtual organisation is a set of co-operating (legally) independent organizations, which to the outside world provide a set of services and functionality as if they were one organisation.

The set of co-operating organizations can change with time; it can be a dynamic configuration depending on the function/service to be provided at that point in time. It can also be a more stable configuration with a sizeable time span and a stable set of services and functions.

The virtual organization paradigm is proposed as a solution to emerging collaboration needs of organizations from variety of domains. Typically, virtual organiza-

Related Content

Understanding Knowledge Bases and Building Membership in Virtual Communities
Susan A. Baim (2006). *Encyclopedia of Virtual Communities and Technologies (pp. 457-461).*
www.irma-international.org/chapter/understanding-knowledge-bases-building-membership/18123

Exploring Virtual Reality for the Assessment and Rehabilitation of Executive Functions
Elisa Pedroli, Silvia Serino, Federica Pallavicini, Pietro Cipressoand Giuseppe Riva (2018). *International Journal of Virtual and Augmented Reality (pp. 32-47).*
www.irma-international.org/article/exploring-virtual-reality-for-the-assessment-and-rehabilitation-of-executive-functions/203066

Fast Single Image Haze Removal Scheme Using Self-Adjusting: Haziness Factor Evaluation
Sangita Royand Sheli Sinha Chaudhuri (2019). *International Journal of Virtual and Augmented Reality (pp. 42-57).*
www.irma-international.org/article/fast-single-image-haze-removal-scheme-using-self-adjusting/228945

Augmented Reality Indoor Navigation Using Handheld Devices
Angelin Gladstonand Aadharshika Duraisamy (2019). *International Journal of Virtual and Augmented Reality (pp. 1-17).*
www.irma-international.org/article/augmented-reality-indoor-navigation-using-handheld-devices/228943

Using a Design Science Research Approach in Human-Computer Interaction (HCI) Project: Experiences, Lessons and Future Directions
Muhammad Nazrul Islam (2017). *International Journal of Virtual and Augmented Reality (pp. 42-59).*
www.irma-international.org/article/using-a-design-science-research-approach-in-human-computer-interaction-hci-project/188480