Federated Information Management for virtual enterprises
Garita Rodriguez, C.O.

Citation for published version (APA):
Chapter 1

Introduction

1.1 The VE Paradigm

In general, most business enterprises nowadays find themselves in a very aggressive environment set by the global market competition and fast commercialization streams. In order to ensure their survival under these hard conditions, enterprises are facing the need to focus on their core competencies and to dynamically establish collaborations with other partner enterprises in order to better react to market opportunities, and to provide better products or services in compliance with strict timing and cost constraints. In fact, this need does not represent a new requirement or trend among enterprises, and it has been repeatedly addressed in the past through the application of different organizational re-structuring processes, such as the outsourcing strategies that emerged some decades ago. However, it is also clear that the continuous evolution of information and communications technology certainly enables the creation of new paradigms and supporting infrastructures aimed at the reinforcement of complex collaborative scenarios among independent enterprises.

In this context, the Virtual Enterprise (VE) paradigm represents a promising solution for those enterprises looking for new information models and technology that allow them to create advanced forms of collaborations with other enterprises in order to adapt to the current market trends. The VE paradigm has also emerged in the last years as an active area of research and technological developments, and it has given place for a large number of projects and initiatives focusing on different facets of this paradigm in different application domains [34, 131, 164, 148, 72]. However, given the extension, complexity and multidimensional nature of the VE concept, there is still no unified or standard conceptual framework and definitions for this paradigm, and a number of different terms are even competing in the literature that refer to similar concepts or to its associated branches. For instance, examples of related terms include: extended enterprises, supply chain management, electronic commerce, enterprise clusters, networked enterprises, and virtual organizations among others; see also [42] for an analysis of some of these related terms.

Although there is no standard definition or common agreement on this concept and
associated terminology, the Virtual Enterprise concept used in this thesis is explicitly stated 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” [42].

Other definitions of the Virtual Enterprise concept can be found in [114, 164, 32, 82, 37]. However, the above definition tries to combine the main elements encompassed by various exiting definitions and terminologies.

According to the described VE conceptualization, the service provision or product manufacturing processes are no longer carried out by a single enterprise; instead, every enterprise is considered as a node in a network of enterprises that adds a given value within a global production cycle (see Figure 1.1). As a result of the collaborative and coordinated interoperation of the individual enterprises, the entire enterprise network operates as a single enterprise towards the accomplishment of a common goal. Therefore, a Virtual Enterprise represents a single entity that is realized through the selected combination of particular skills and resources from different companies. The benefits and other implications of forming this kind of enterprise alliances have also been analyzed in [52, 61, 81, 94, 57]. Furthermore, it is important to bear in mind that every enterprise can simultaneously act as a member of different VEs and play different partner roles in each of them.

Figure 1.1: Example of collaborating manufacturing Virtual Enterprises.
1.1. The VE Paradigm

Although some support for enterprise cooperation has been already provided in some areas for instance, focused on the technical product design and engineering in the automobile industry, the main issues addressed by the VE paradigm concentrate on the high level of integration among pre-existing heterogeneous companies, and on the proper "harmonization" of selected information and communications technology. In other words, the provision of the support for the entire life-cycle of such virtual entities consisting of autonomous heterogeneous enterprises that need to act as a single unit, represents a challenge that needs to be entirely addressed by existing advanced technology and computational infrastructures. Furthermore, in addition to the technological components and infrastructures necessary to materialize the VE paradigm, there is another unresolved challenge, which is to define and develop a general reference architecture for the VE support platform, its associated tools, and the protocols and mechanisms that would enable the desired level of cooperation among a large number of potential partner enterprises around the world.

In relation to the existing information technology that can be applied in the design and implementation of VE support infrastructures, there is an extremely wide variety of approaches, tools, components, models and standards that can be considered [41, 70]. The following general areas need to be evaluated for their application in the VE infrastructure: distributed/federated information management systems, workflow management techniques, standard information models and protocols (e.g. EDI, STEP), ontology management systems, advanced Web applications technology, secure and reliable communication protocols, high-performance computing environments, distributed object management architectures, and multi-agent systems among others. One issue that has however become clear within the VE research community, is that the materialization of the VE paradigm requires a level of organization and adaptation that cannot be totally supported by traditional plug-and-play portals, that are sufficient for simple service and product provision scenarios through the Web. Furthermore, once the proper technological components and approaches have been evaluated and selected to be applied into a given VE platform, it is necessary to support their integration and interoperability.

Some of the technological issues described above, are currently being tackled by several VE projects and initiatives aiming at the support for different aspects of the VE paradigm, addressing various application domains. These initiatives involve international R&D projects, university research efforts, and industrial initiatives among other kinds of collaborations. The main projects more closely related to the work presented in this dissertation include: NIIP, VEGA, X-CITTIC, PerDis, PRODNET, MASSYVE, FETISH, COWORK, LOGSME, EisNet, GEN, and VIRTEC. These projects will be analyzed in details in Chapter 2 of this thesis.

Besides the obvious impact of the VE paradigm on the technological infrastructure of existing companies, the new forms of inter-enterprise collaborations induced by this paradigm will also drastically change the existing cultural, legal and organizational settings of the member enterprises. For example, the involvement in VEs may require the introduction of new work-related concepts that need to be handled by individual employees working for different companies and interacting with the VE platform, including different emerging authority models, shared responsibility sense, protection
of information privacy, etc. Namely, the VE concept also represents a challenge for industrial sociology and international commerce legislation regarding the definition and reinforcement of a new unification framework between people, organizations and society, supported by the forthcoming technology [41, 112, 165, 108].

1.1.1 VE Characterization

Due to the multiplicity of organizational variants, Virtual Enterprises can be classified through many different perspectives [44, 42]. The fundamental characteristics describing the VEs analyzed in this section include:

**Alliance duration:** a VE can be established in order to satisfy a single business opportunity, and when that objective is accomplished the alliance dissolves. On the other hand, the duration of an alliance may have a specified time span, or continue for an indefinite number of business processes.

**Topology variation:** the structure or topology of a certain VE in terms of partners and their particular roles can remain fixed until the target VE goals have been achieved, or it can be dynamically transformed due to for instance, resignation of participating partners, joining of new enterprises, changes in the VE coordination or production roles assumed by different members, etc. Furthermore, depending on the kind of VE being addressed, it may be necessary to allow partners to dynamically join/leave the VE.

**Participation exclusivity:** this characteristic determines the possibility of involvement of an enterprise in only a single or multiple VEs. If enterprises can be simultaneously involved in several VEs, they must be able to distinguish which information must be made available to every particular VE partner in any given VE. For instance, in case of two enterprises taking part simultaneously in two different VEs, disjoint sets of local data may be shared among the same enterprises, within each VE in which they participate.

**Coordination approach:** the approach taken by VEs in order to coordinate and follow up the progress of the tasks being carried out by VE members can be quite varied. For instance, in some business sectors such as automotive industry and agri-business, the establishment of direct supplier/clients interactions among partners frequently occurs. In these cases, there is typically one company that is “served” by a relatively fixed network of suppliers, which constitutes a star-like coordination structure. While in other situations, a supply or service chain may be defined without the existence of a dominant company, where enterprises can cooperate in a more democratic manner. Here, the coordination structure is usually defined through the joint agreement of the key enterprises within a “democratic alliance”. In general, other forms of coordination may eventually emerge and be accepted among VE member enterprises [18].

**Visibility scope:** this characteristic refers to the extent in which a given VE member is able to access and retrieve information from other partners which can be either
directly or distantly related to it, in terms of the given production or supply chain relationships. In other words, the visibility scope determines how far a given partner enterprise can “see” inside the graph represented by the VE network, even beyond the scope of the immediate VE collaborating partners. For instance, the enterprise at the beginning of a sequential production/supply chain may need to monitor activities being carried out by another enterprise that is positioned completely at the end of the chain (i.e. several nodes apart in the production chain). Deep visibility levels are necessary for proper coordination and supervision of the production activities within certain VEs. Such a multi-level visibility scope is also necessary to implement “demand forecast” functions based on information gathered from different members in the supply chain [42].

1.1.2 VE Life Cycle

A Virtual Enterprise represents a complex and dynamic entity that undergoes a sequence of stages during its life cycle. Currently, there is no common agreement on the definition of a reference VE life cycle, and several different life-cycle models have been used depending on the characteristics of the VE application domain and scenarios being addressed [135, 136, 121]. In this section, the minimum VE life cycle that is used as a reference in this thesis is presented (for a more detailed description, please see [42, 40]). Figure 1.2 shows the basic VE life cycle phases, as well as its preceding and succeeding stages. Please notice that a given enterprise may be involved in several VEs, each going through different life-cycle phases at the same time. The main VE life-cycle phases and associated steps are briefly described below:

1. Pre-VE cycle stage. This stage takes place before the creation of any Virtual Enterprise and basically considers the steps that are required by an enterprise in order to become a node in the network of eligible VE partners. During this stage, an enterprise will install and configure the software layer that will allow

![Diagram of VE life cycle stages](image-url)

Figure 1.2: Main Virtual Enterprise life cycle stages and associated steps.
it to get involved in actual VEs. This stage involves two main steps:

- Setup of Infrastructure Resources: every enterprise must determine and acquire the computational resources that are required in order to become a potential "VE member enterprise". Thus, this point represents a preparatory step before installing and configuring the required software that will enable the operation of the company as a VE member. This step may include for example, the purchase of computational equipment, software (e.g. database management system, operating system) and Internet connection services.

- VE node adaptation and manifestation: once the existing enterprise has properly installed the resources specified in the previous step, it needs to install a certain software to serve as its VE support layer, referred to as "VE Cooperation Layer" (VCL), which would actually enable it to properly act as a VE member in future collaborations. This layer is represented as the gray area in the enterprise nodes in Figure 1.2. Please notice that as mentioned before in this chapter, the architecture of the VE platform cannot be entirely based on typical "portal" approaches, in which a broad array of resources and services, such as e-mail, forums, search engines, on-line shopping malls, etc., are offered through an enterprise Web site. Namely, the VE paradigm requires a much higher level of cooperation and interoperation among enterprises and their internal systems, which is not well supported by portals and their offered services. Thus, the need of a "layer" for VE support at each enterprise node is identified here, in order to handle the complexity of both cooperation and interoperation issues among VE member enterprises. The adjustment of the VCL to comply with the existing enterprise procedures and to interact with internal modules is performed at this stage. After the VE Cooperation Layer is installed, the new node would be announced within the network of potential VE members, so that it can be eventually considered as a partner by other nodes. For this "manifestation" step, the enterprise can be for instance registered in a public directory including its general profile and other information that may be interesting for other potential partners.

2. VE Creation. This is the phase in which the VE is created and configured. This phase involves the following steps:

- Identification of business opportunity and definition of VE goals. In this step, a business opportunity is identified, and a VE is conceived by a given "initiator" enterprise, in order to fulfill the target objectives. Here, the initiator enterprise identifies the main resources or competencies that can be provided by other companies within a possible VE consortium.

- Partners search and selection. Once the general VE goals are established, the suitable VE partners must be rapidly searched, identified and selected. These partners can be selected from internal lists of previous partners,
or from other sources of information such as for instance public/private partner directories and Internet services offered by existing companies [46]. If necessary, call for tenders can be generated in order to collect bids with more specific information from potential partners. Furthermore, this step may be supported by VE brokers, which enable the creation of VEs through the use of services provided by global industry clusters [31]. Please notice that the partner search and selection step is necessary not only during the creation phase of the VE, but also during the operation phase in case a given VE partner leaves the VE for any special reason, and needs to be replaced by another enterprise.

- **VE initial agreement and topology definition.** In this step, the general VE structure and topology is outlined and initially agreed among the VE partners. To reach this initial agreement, a corresponding negotiation process needs to take place. Furthermore, the VE partners may also agree on some common ontology definitions and information management standards to support their future interoperability requirements.

- **Contract negotiation.** After the VE has been initially agreed and defined, a “VE contract” must be negotiated in order to clearly establish the relationships and responsibilities among the VE members. This step may involve the definition and distribution of some “supervision clauses” that may determine the level of visibility on information access among VE partners, as well as the reporting duties of partners, among other aspects. In general, the proper establishment, representation and management of contracts that regulate the relationships among VE members, is a complex issue that still requires further research (see also [88, 121]).

- **VE partners configuration.** Here, each VE partner defines and configures its local VCL in order to act in accordance with both the internal company procedures as well as the agreed VE topology and contract. This step is done specifically for each VE in contrast to the general VCL configuration that needs to be done during the pre-VE life-cycle stage. For instance, contract clauses and/or VE network definitions can be loaded at every VE member in order to define the access rights on local information for every “other” VE partner. Furthermore, workflow-based definitions of the VCL behavior and certain communication settings may also need to be configured for each specific VE in which a given partner is involved. All this information may need to be re-configured during the VE operation and evolution phase described below.

3. **VE Operation and Evolution.** This phase mainly embodies the execution and coordination of the processes that need to be carried out towards the fulfillment of the VE goals. Furthermore, specific steps need to be considered in order to support the dynamic evolution of the VE entity. The steps involved in this phase include:

- **Basic information exchange among VE partners.** This point is necessary
in order to support for instance the basic collaboration among the VE members. It can include the exchange of: generic partner messages, commercial order data, technical product information, delivery reports, VE configuration data, etc.

- VE coordination and monitoring. The monitoring and coordination of the VE tasks and activities is a crucial issue during the operation phase of the VE life cycle. Given the fact that the VE sets forward one or more global goals involving several partners, which in principle can be completely autonomous and independent, there must be a well-defined mechanism to coordinate and monitor their inter-related activities towards the achievement of these global goals. Specific partners within a VE may assume different coordination and monitoring responsibilities, as will be explained in Section 1.1.3.

- VE evolution. This point refers to modifications in the general VE structure or topology due to the departure, replacement, changes in roles, or incorporation of new VE members. This step implies the reconfiguration of the VE cooperation layers to reflect the new VE structure and relationships among the members. Due to changes in partners memberships and roles, new contracts may need to be negotiated among the partners to reflect the evolution of the VE.

- Events and exception handling. Certain asynchronous events and exceptions that can occur during the VE operation, need to be properly handled, in order to guarantee the consistency of the global VE network. These events can be generated either internally within the company or from other VE nodes. For example, the generated exceptions may be related to internal production delays, changes in product delivery dates, communication channels disruptions and emergency situations, among others.

4. VE Dissolution. This stage is reached when the VE operation and evolution phase finishes either because the final goals were achieved, or because the involved partners decided to cease their allied cooperation. At this point, proper actions must be taken in order to dissolve the VE according to the legal and contractual agreements defined in previous stages. This phase involves for example: the assessment of accomplished VE goals, evaluation of partner obligations and eventual liabilities, “disassembly” of the VE structure and partner relationships, reconfiguration of access rights by each partner, and gathering of partner performance and historical information to be used when creating future VEs.

5. Post VE-life-cycle stage. The VE is dissolved after the VE life cycle is completed and the collaboration among the involved enterprises is finished, although some of the enterprises may still remain partners and collaborate in the context of other existing or future VEs. In this post-VE stage, valuable information may be gathered by looking backwards at the results of the different stages of VE life cycle. For instance, knowledge about partners performance/reliability, jointly designed business processes, successful (and unsuccessful) VE scenario cases,
etc., may provide partners with useful feedback to be applied in future collaborations.

Finally, the support for agility during all the stages of the VE life cycle described above, represents a particularly important requirement for the successful operation of the VE [45]. Namely, the VE infrastructure should provide the required mechanisms and tools in order to detect and rapidly react and cope with unpredicted environmental changes during the all the life cycle stages (see [82, 98]).

1.1.3 VE Member Roles

Besides the evident role that a VE member can play in relation to for instance, a production chain in a manufacturing VE, e.g. supplier, client, distributor, etc., certain enterprises need to play roles that are specifically related to the generic VE structure and operation. Examples of some VE member roles include: VE Member Enterprise, VE Coordinator, and Network Directory Node (see also [42, 44]). These roles are described in more details next:

a. VE Member Enterprise. Every enterprise in the network of nodes that have been enabled with a VE Cooperation Layer may eventually take part as a member in a VE. The main functionalities associated with a VE Member Enterprise include:

- Execution of the VE tasks that have been assigned to this VE Member Node according to the VE contract or general agreements.
- Establishment of contact and interoperation with other VE nodes when necessary.
- Sharing and exchange of part of its local information with other VE Member Nodes in the network.
- Sharing and exchange of specific status information required for global control and monitoring of VE tasks.
- Definition of the proper access rights and visibility levels on local information that needs to be shared with other VE nodes.

b. VE Coordinator node. An enterprise in the VE network can take the coordination role in a VE. As a coordinator, this VE node may provide the following functionalities (some of these tasks may be shared or delegated to several other nodes with subordinated roles, e.g. VE supervisor nodes):

- Configuration/reconfiguration of VE topology.
- Distribution of the VE contracts and supervision clauses information among VE partners.
- Periodic control and monitoring of tasks assigned to VE Member Enterprises.
• Conflict detection during VE operation and analysis of solution approaches.

c. Network Directory node. One or more nodes in the network of actual and potential VE Member Enterprises may act as the network directory nodes. Directory nodes may operate independently of each other, or they may work together to maintain a coherent set of information. Obviously in the second case, for any node in the network it is enough to have access to one directory node only. The main tasks of a directory node include:

• Provide information on the enterprises in the network. Thus, enterprise profiles, mailing lists, classifications on product/services, etc. can be stored and maintained. This information is very important in order to support the partner search and selection steps during the VE creation phase.
• Provide a public “board” for announcements that may be needed to be used, for instance for partner search in the network.

Finally, please notice that VE network nodes may also assume other roles (e.g. VE brokers [31]), depending on the conceptualization of the VE paradigm being used.

1.2 Thesis Objectives and Proposed Approach

As mentioned previously in Section 1.1, although the currently available information and communications technology resources and tools provide solutions to certain specific technical problems that arise when supporting certain basic interactions among enterprises, there are still many obstacles and open issues that need to be properly addressed when supporting complex collaborations among enterprises involved in VEs.

In particular, one of the most challenging issues in the support of Virtual Enterprise platforms is the sharing and exchange of information with proper mechanisms and regulations among pre-existing heterogeneous and autonomous enterprises and their internal systems. It is clear that without an adequate support framework for information management, it is impossible for enterprises to collaborate as a single virtual entity.

Among the key problems faced for information management approaches supporting the VE domain we can mention:

• Lack of standard definitions of information models and access mechanisms.
• Support for sharing and exchange of distributed information, while maintaining the proper level of autonomy and security for each VE member.
• High degree of heterogeneity encountered at every VE node.
• Wide diversity of applicable information technologies and tools.
• Functional generality and extensibility.
• Adequate performance and scalability.
Therefore, the design and implementation of a distributed information management system aimed at the support of VE infrastructures must address these general challenges, as well as many other specific requirements related to the particular application domain under consideration.

1.2.1 Thesis Objectives

In order to address these information management challenges for the support of Virtual Enterprise infrastructures, the general objective of this thesis is:

"The analysis, design and implementation of a federated/Distributed Information Management System (DIMS), specifically tailored to properly support the complex requirements set forward by Virtual Enterprise collaborative scenarios."

The DIMS will be responsible for representing, managing, and providing the proper means for accessing the information which is necessary for the operation of all VE Cooperation Layer components. The general interactions in terms of exchange of data and control information between two enterprise nodes is shown in Figure 1.3.

This figure shows the general data and control information flow among two enterprise nodes, from the DIMS point of view. Namely, the DIMS will provide the required functionality to support the proper information management of the data coming from: the local internal enterprise management systems, the other VE Cooperation Layer components, and also other DIMS components located in remote VE Cooperation Layers. Since a good part of this information is distributed, one of the major tasks of the DIMS is to properly support its features, with the expected data location transparency for the user, and the necessary transparency for distributed query processing, site autonomy, and reliability, among other requirements.

Furthermore, other specific objectives associated with the development of the federated DIMS include:

![Diagram of Information Flow](Figure 1.3: Role of the Distributed Information Management System for VE support.)
• Support general collaboration scenarios among VE member nodes through the provision of specialized VE-oriented data management services.

• Support for integration of the information that is made available from one VE node into another one. As a result, the DIMS will support the seamless access to up-to-date information that is physically distributed among the VE nodes.

• Reinforcement of the concept of enterprise autonomy, where enterprises have their own information, and decide by themselves which subset of this information is made available to other determined members.

• Definition of an information access mechanism among the VE nodes, through which secured fine-grained access rights and visibility levels can be defined locally at every node to determine which other VE partners are allowed to access which part of the local information.

• Support for storage and management of information handled by each component module of the VE Cooperation Layer.

• Proper support for handling information represented according to specific standards and/or data models within the VE Cooperation Layer e.g. EDI, STEP. Furthermore, interoperability issues among these standards must be addressed when necessary.

• Storage and management of a wide variety and kinds of information associated with the internal business processes defined at each enterprise.

• Provision of an interoperability mechanism to support data exchange functionalities with internal enterprise systems.

• Provision of data query/update operations through diverse mechanisms, including for instance SQL queries, generic data management functions, high-level functions for specific components of the VE layer, and generic distributed queries.

• Development of flexible and configurable information management functionalities in order to support the evolution of the VEs as they follow the different stages of their life cycles.

• Use of secure and reliable network communication services and protocols in order to guarantee the protection of private enterprise information.

In the next section, a summary of the general approach that has been followed in order to achieve these objectives, is presented.
1.2.2 Summary of General Solution Approach

The wide range of complex characteristics describing the VE paradigm introduce many information management requirements that need to be addressed by an IT platform aimed at the support of this form of enterprise collaborations. A set of general challenges faced in the design and development of the information management system of a given VE support platform, were already introduced in Section 1.2. Many of these requirements stem from the high degree of heterogeneity and autonomy of the pre-existing nodes that simultaneously need to act together as a single unit. However, some of these general requirements may become more relevant or acquire a different significance depending on the VE application domain and the target objectives of the particular VE support platform under consideration.

Therefore, as a first step towards the accomplishment of the objectives stated in Section 1.2.1, an extensive analysis of the information management requirements for the DIMS was carried out considering different scenarios to support different stages of the VE life-cycle, with special focus on Small- and Medium-sized Enterprises (SMEs) in the industrial manufacturing sector.

Based on this extensive analysis of the information management requirements for DIMS, one main conclusion is that to handle the complex requirements set by the VE paradigm, an advanced federated/distributed information access mechanism among the VE nodes is required. In order to briefly introduce the motivation for applying federated information management in VEs, we need to consider the fact that Virtual Enterprise members need to have access to up-to-date information that is physically distributed among different nodes in the VE network. This is necessary in order to support for instance, the basic data exchange operations among VE members, as well as more sophisticated functionalities regarding the coordination and monitoring of the tasks that are being independently executed by different VE members. Clearly, an advanced information management mechanism must be designed and implemented in order to provide seamless access to VE distributed information.

Furthermore, even though enterprises involved in a VE must share and exchange a part of their information in order to achieve the common VE goal, it is also true that not all members of a VE play the same role and not all of them should have the same access level to the information stored in other enterprises. It is clear that among competitive enterprises in a VE the amount of trust is limited, and that every enterprise needs to precisely define the specific access rights and visibility levels on its information for every other VE partner. As a result, within the VE, support for the security and provision of different rights to access shared data – mostly based on other enterprise’s role in the VE – are required to be provided and reinforced.

In order to address these and many other issues identified during the requirement analysis phase, the approach for VE information management adopted in the design of the DIMS is based on a federated database architecture [71, 11, 65, 74]. Through this federated approach, data can be imported/exported/integrated among VE members, but the proper access rights are defined locally at every enterprise to precisely specify the rights of external nodes. The federated architecture approach has proven to adequately facilitate and support the sharing and exchange of distributed information.
between enterprises in VEs, while providing the necessary information visibility levels to ensure their own autonomy and information privacy. The design of the DIMS federated architecture has its roots in the PEEER federated database system [169, 16].

In summary, the general design of the DIMS federated architecture is defined by several elements or components, among which the following can be briefly mentioned in this section: the VCL Integrated Schema, the Export Schema Manager, and the Federated Query Processing component. The functionality of these and other internal DIMS components will be described in details in Chapter 4 of this thesis.

The practical application of the DIMS architecture has been validated in the context of several real-case VE demonstration scenarios. Namely, the main features of the DIMS have been specifically applied to support several steps of the VE life-cycle described in Section 1.1.2, including for instance, the protected exchange of distributed information among regular partners, and the VE coordination and monitoring tasks that take place during the VE operation phase.

All the DIMS components and functionalities presented in this section were fully implemented at the University of Amsterdam, and were properly validated, tested and integrated within the context of the VE Cooperation Layer architecture designed for the PRODNET II project. This project is described in details in the next section.

1.3 Related Projects and Scientific Publications

This section describes several research and development projects and scientific publications that have directly or indirectly contributed to this thesis work.

1.3.1 PRODNET II Project Overview

A summarized description of the PRODNET reference architecture is provided in this section, since this project served as the main context in which the framework for federated information management for VE support presented in this thesis, has been developed. For a more detailed description of this project, please see [40, 39].

The European ESPRIT project PRODNET II (1996-1999) aimed at the design and development of an open IT platform to support industrial virtual enterprises, with special focus on the needs of Small- and Medium-sized Enterprises (SMEs).

One of the general requirements identified in PRODNET in order to develop a VE support infrastructure was the fact that the involved SMEs must be able to inter-operate and exchange information in real time so that they can work as a single integrated unit, while at the same time keeping their own independence and autonomy. Furthermore, these enterprises typically already exist before they decide to join in an information sharing and exchange network. Consequently, every enterprise is autonomous, developed independently of other enterprises and uses the specific information management and control strategies that serve its purposes best. The situation is thus one of great heterogeneity and requires the adaptation of existing Production Planning and Control (PPC) systems to a much more open and distributed virtual production environment.
To support this VE environment and to properly cope with legacy systems in SMEs, PRODNET II proposed an infrastructure considering three main components: the PRODNET Cooperation Layer (PCL), the Advanced Coordination Functionalities (ACF), and the Internal Module (see Figure 1.4).

The Internal Module of a node basically consists of the internal Enterprise Resource Planning (ERP) systems of the company, such as its Production Planning and Control (ERP/PPC) system, as well as other engineering systems necessary to accomplish its regular operations [49]. The functionalities supported by the ERP/PPC system include for example: (a) Industrial Logistics Management: Orders flow management, Product data management, Sales Forecasts handling, Actual Requirements Planning; (b) Master Production Scheduling; (c) Production Control; (d) Quality Control / Tracking; and (e) Industrial Costing.

The PCL component is responsible for the actual inter-operation between a given node and other nodes in the network. PCL provides a wide variety of services that allow the enterprise to interoperate with others in the context of the VE. The PCL itself consists of several internal components:

- **LCM - Local Coordination Module.** This component is responsible for implementing the desired behavior as defined in the configuration phase of the VE [47]. It handles all the cooperation events related to a given VE member enterprise according to the specified rules for the particular enterprise. These events have an asynchronous nature and are generated either by other nodes of the VE, by the Internal Module of the enterprise, or by the Human Interface. LCM acts as a workflow engine according to the reference model of the Workflow Management Coalition [167].

- **DIMS - Distributed Information Management System.** The DIMS of the PRODNET Cooperation Layer is responsible for modeling and managing the exchange of all integrated VE cooperation-related information, while preserving the autonomy and information privacy of the involved enterprises [71, 65, 74]. Different levels of information visibility for other enterprises, defined at every enterprise,
guarantee the enterprise's autonomy and sharing of information. Furthermore, information privacy and ownership are also preserved when an enterprise is involved in more than one VE. The PRODNET DIMS represents the actual system in which the federated information management approach for VE support proposed in this thesis, has been applied and validated.

- **EDI Module.** This module is responsible for receiving and formatting order-related messages in EDIFACT format [79]. Among other functionalities, the EDI module parses EDIFACT messages (for various versions of the standard), checks for completeness of message contents, and generates appropriate formats for sending out EDI messages. It also detects and extracts information embedded in other kinds of EDI messages such as order-associated STEP specifications (via EDIFACT CONDRA messages).

- **STEP Module.** The major task of the STEP module is to handle the technical product data used within PRODNET [142]. Ideally, all product data should be exchanged in STEP format. The STEP services provided in PRODNET allow the transmission and reception of STEP files between two VE member companies. Complementarily, higher-level functionalities for product data visualization and product administrative data management are provided.

- **PCI - PRODNET Communication Infrastructure.** This module is responsible for handling communications with the other nodes in the network [119]. It includes functionalities such as: selection of communications protocol and channels; basic communications management; privacy reinforcement mechanisms; and secure message communication channels between nodes. Regarding the security aspects, the following features are included in PCI: symmetric and asymmetric cryptography, authentication based on signed certificates, adoption of emerging security standards, digital signature and certificate management, and tunable security levels.

- **Configuration and User Interface.** The PRODNET platform is intended to support a large diversity of enterprises and interconnection modes. Therefore, it is necessary to specify the desired cooperation behavior, regarding both the business processes and the information exchange, in terms of an explicit plan (each enterprise has to define its particular behavior) that will be "executed/controlled" by the Local Coordination Module (LCM). Additionally, the Configuration Component allows a specification of the structure of the VE and the access rights of all its members. The User Interface offers a friendly interface between the human operator and the PCL. The level of human intervention in this process will depend on the policy of each company and will be specified at the configuration step of the VE creation phase through the configuration specification and the workflow plan definition.

Furthermore, the Advanced Coordination Functionalities (ACF) of PRODNET has been conceived in order to address some advanced VE life-cycle functionalities such as: i) Partners search and selection; ii) Integrated logistics decision support,
1.3. Related Projects and Scientific Publications

which aims at providing the enterprise with “real-time” information about the ongoing operations of its current suppliers. The ACF integrated logistics support has been implemented as a system called “Distributed Business Process Management System” (DBPMS). The main objectives of the DBMPS are to provide reliable and periodic information about the supply-chain, and to support rapid decision making, in order to improve the agility of the enterprise and hence its competitiveness [99].

It can also be mentioned that a complex demonstration scenario was developed in order to demonstrate and assess the architecture and infrastructure proposed by PRODNET [43]. This demonstration took place at the Pro-VE’99 conference in October 1999 in Porto, Portugal. The reference scenario involved the functionality of all the aforementioned PCL modules in order to support the creation and operation phases of Virtual Enterprises composed of European and Brazilian enterprises. The demonstration scenario was formulated based on real business requirements identified by actual end-user companies.

The PRODNET II project was partially funded by the European Commission (ESPRIT programme), and by the Brazilian research council (CNPq). The PRODNET project consortium was represented by three kinds of European and Latin American partners: software development companies, universities and research institutes, and end-user SMEs. The specific partners are grouped according to this classification in Figure 1.5.

1.3.2 Other Related Projects

Besides the PRODNET II project, which served as the main context to develop the VE federated information management approach presented in this thesis, there have been several other projects that have directly or indirectly contributed to, as well as

<table>
<thead>
<tr>
<th>Software Development Companies</th>
<th>Universities / Research Institutes</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSIN - Construção de Software e Automação na Indústria Lda. - Portugal</td>
<td>UNL - Universidade Nova de Lisboa Portugal</td>
</tr>
<tr>
<td>ESTEC - Estudos e Tecnologias da Informação Lda. - Portugal</td>
<td>UvA - Universiteit van Amsterdam The Netherlands</td>
</tr>
<tr>
<td>LICHEN Informatique - France</td>
<td>UFSC - Federal University of Santa Catarina Brazil</td>
</tr>
<tr>
<td>ProSTEP GmHb - Germany</td>
<td>UNINOVA - Instituto de Desenvolvimento de Novas Tecnologias - Portugal</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>End-user SMEs</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>MIRALAGO S.A. - Portugal</td>
<td>HERTEN - Engenharia de Moldes Brazil</td>
</tr>
</tbody>
</table>

Figure 1.5: The PRODNET II consortium.
benefited from the research work developed in this thesis. For instance, the FETISH project has allowed the extension and validation of the proposed federated information management approach in another VE application domain, i.e. the tourism Virtual Enterprises. Furthermore, several other projects in which the author has been involved as member of the Cooperative Information Management (CO-IM) database group of the University of Amsterdam (UvA) [55], have indirectly contributed to this dissertation work in the sense that they have provided several reference scenarios for which different aspects of the distributed and federated information management approach have been developed and applied. Namely, through the partial involvement in these projects, the author has gained valuable experience and technical skills that have been implicitly applied to achieve the work presented in this dissertation. A brief description of these projects is provided below:

- **European 5FP FETISH project 13015 – Federated European Tourism Information System Harmonization (2000-2002).** The general goal of FETISH is to integrate the fragmented tourism information systems and their IT-based services into a federation of distributed resources that are presented through a single infrastructure to end users and other service provider enterprises [8]. Furthermore, the project also aims at supporting the concept of Virtual Enterprises in the tourism sector, based on the VE framework developed in PRODNET II. In FETISH, the VE paradigm is applied in order to promote and reinforce the proper cooperation among service provider enterprises that can work together in order to offer new high-level value-added services, which are in turn defined as a composition of other existing basic and/or value-added services. The main task of the CO-IM group of UvA in this project concerns the development of the federated/distributed information management component that can properly support the requirements set by the FETISH VE application domain. The author has been actively involved in the requirement analysis, system design, and development phases of this federated information management system.

- **Dutch ICES-KIS HPCN project Virtual Laboratory (1999-2003).** The Virtual Laboratory (VL) project initiated at the University of Amsterdam aims at the development of a hardware and software reference architecture, and an open, flexible and configurable laboratory framework to enable scientists and engineers to work on their experimentation problems by making optimum use of modern information technology approaches. The general design of the VL architecture is based on multiple functionality layers, so that the application- and domain-specific computational and engineering issues can be separated from the generic computing aspects. In particular, the VIMCO co-operative information management layer, which constitutes the main task of the CO-IM database group of UvA, provides archiving services as well as the information handling and data manipulation within the virtual laboratory. This layer supports a wide range of functionality ranging from the basic storage and retrieval of information (e.g. for the raw data and processed results) to advanced requirements for distributed information integration. In this context, the author has been
1.3. Related Projects and Scientific Publications

partially involved in the requirement analysis and system design aspects of the internal components of the VIMCO layer [6, 13].

- **European ESPRIT IV Waternet project 22.186 (1996-1998).** This project aimed at the design and development of a knowledge capture and knowledge management system towards the control, optimal operation, and decision support for water distribution networks. The Waternet system developed in this project, assists the distributed control within a water management network through the utilization of several subsystems in charge of the following functionalities: machine learning, optimization, simulation, water quality monitoring, federated information management, and the supervision system. In order to support the cooperation and the sharing and exchange of information among remote water management nodes, each node in the network has been extended with a federated cooperation-layer, supported by the federated information management subsystem. Please notice that although the application domain of this project is not related to Virtual Enterprises, it does share many similar characteristics associated to a network of autonomous, distributed, and heterogeneous nodes. In this project, the research developed by the CO-IM database group of UvA focused on the analysis, design, and implementation of this federated information management component based on the PEER federated database system. The author participated in the early requirement analysis and design phases of this federated database system [4].

- **Dutch HPCN project 3DOME (1996-1998).** This project (in cooperation with several Dutch industries and software houses) aimed at the design and development of a 3D graphic object mediator center on the Internet. The mediator center provides a graphic object-broker facility (the model shop), high-performance 3D model render services (the render farm), and additional complementary services, such as a talent pool and graphic software tools for rent. In this project, the CO-IM database group of UvA developed the information management system required for all 3DOME broker services. This system handles a library of hierarchically defined objects and a catalog of the 3D models accessible through the Internet. Different kinds of access, through the diverse views (e.g. browsing and administration) defined on the graphic data needs to be supported. Due to the brokerage nature of this application, different security levels for access to database, and efficient database search mechanisms are required. Some of the information management addressed in 3DOME can also be found in certain scenarios related to the application domain of VEs (e.g. partner profile and product/service catalogue management). In this project, the author assisted with the design and development tasks associated with the information brokerage system.

1.3.3 Summary of Related Scientific Publications

Most of the content of this thesis has been published earlier in different forms including several book chapters, journal articles, international conference papers, newsletters
contribute, and technical reports. Below, an executive summary is provided in Table 1.1, regarding the publications of the author, according to different subjects addressed in this thesis. The complete list of publications achieved by the author is provided in Appendix A of this thesis.

<table>
<thead>
<tr>
<th></th>
<th>Book Chapters</th>
<th>Journals</th>
<th>Conferences</th>
<th>Newsletters</th>
<th>Technical Reports</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>General VE Paradigm Issues</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Survey of VE Information Management Approaches</td>
<td>1</td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>DIMS Requirement Analysis</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>DIMS Design and Implementation</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td></td>
<td>3</td>
<td>10</td>
</tr>
<tr>
<td>DIMS Extension and Applications</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>Related Virtual Collaborative Environments</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Totals</td>
<td>5</td>
<td>3</td>
<td>7</td>
<td>1</td>
<td>7</td>
<td>23</td>
</tr>
</tbody>
</table>

Table 1.1: Executive summary of author's publications.

Table 1.1 indicates that a relatively high amount of publications have been produced specifically addressing issues related to the VE paradigm, VE distributed information management, and other related virtual collaborative environments (e.g. Virtual Laboratories).

1.4 Structure of Thesis Document

This thesis document is organized in several chapters as described next.

Chapter 1 defines and characterizes the VE paradigm as the general target domain considered in the thesis. Furthermore, this chapter also provides an overview of the thesis research work in terms of the main challenges to be faced, the target objectives, and the general solution approach that has been taken in order to properly address the distributed information management requirements for a generic VE support infrastructure.

Chapter 2 presents an extensive analysis of distributed information management approaches for Virtual Enterprise infrastructures, that are closely related to the research in this thesis. In order to accomplish this goal, this chapter first provides a survey of related information management technologies, including general approaches to manage distributed information, as well as some relevant information management standards and tools. Secondly, this chapter presents a study of information management approaches that have been applied in several actual VE-support research and development projects. These projects represent a selected group of relevant VE support infrastructures, which are described and compared according to a predefined set of information management characteristics specially designed and introduced in Chapter 2 for this evaluation purpose.
Chapter 3 focuses on the analysis of the specific VE information management requirements for the Distributed Information Management System (DIMS) component of a generic VE support infrastructure. These requirements include both the information modeling and the functional requirements for the DIMS. Furthermore, based on the results of the performed analysis, this chapter proposes the federated information management architecture as the support framework for effective information sharing among the VE member enterprises.

Chapter 4 aims at describing the general design and implementation of the DIMS federated architecture and its major elements, including for example the integrated federated schema, the Export Schema Manager, the Federated Query Processor, and the internal DIMS kernel. In addition, Chapter 4 also illustrates the combined application of these components and mechanisms in the context of real application cases from the manufacturing sector, used for the VE demonstration scenario.

In Chapter 5, the main objective is to demonstrate and validate how the general federated information management approach that has been presented in Chapters 3 and 4, can be tailored and extended in order to cope with the requirements found in different VE application domains, taking as an example the tourism service provision sector. For this purpose, the requirement analysis, general design and development phases for a federated information management system supporting service-oriented Virtual Enterprises in the tourism application domain, are presented in this chapter.

Finally, Chapter 6 summarizes the main conclusions derived from this research and provides some directions for further potential extensions and future work.