Semantic Web and Peer-to-Peer: Integration and Interoperability in Industry

Vagan Terziyan, Andriy Zharko

Industrial Ontologies Group, MIT Department, University of Jyväskylä,
P.O. Box 35 (Agora), FIN-40014 Jyväskylä, Finland
vagan@it.jyu.fi, anzharko@cc.jyu.fi

ABSTRACT

Companies need integrated information systems to react to requirements of their customers, collaborate with partners, identify and exploit new opportunities quickly and effectively. A need for solid technology that will provide effective Knowledge Management within and across industrial enterprises is crucial. Nowadays a new integration concept based on Semantic Web is being developed, which pretends to address the industrial challenges concerning integration and interoperability. The concept assumes semantic annotation of Web resources and services to make them "understandable" by computers. On the other hand, emerging Peer-to-Peer technology and appropriate networks suite well to the increasingly decentralized nature of modern companies and their industrial and business processes, whether it is a single enterprise or a group of companies. In this paper a potential of combination of Semantic Web and Peer-to-Peer is analyzed in context of integration and interoperability of industry enterprises. To distinguish a target a brief analysis of an industrial information infrastructure is made. A survey of existent solutions is made that gain benefits from Semantic Web and Peer-to-Peer to provide evidence in favour of these two technologies combination. The real industrial case of the global network of maintenance services is described that implements the above concepts within the *OntoServ.Net* project.

Keywords: Semantic Web, Peer-to-Peer, Agent Technologies, Interoperability, Intelligent Web Services, Smart Devices.

1. INTRODUCTION

To react fast to unique needs of each customer companies should compose their products from interchangeable modules. Such approach makes products highly configurable (Siltanen et. al., 2001). New products are based on standard components used in earlier deliveries and new component variations. Furthermore such components may be from different manufacturers. Provision of interoperability between heterogeneous decomposed modules is extremely important (Automation, 2002a).

After starting the manufacturing process, an efficient discontinuities management (Pyötsiä, 2002) must be provided. This is aimed to provide the permanent run of a production system without shutdown to fulfil orders of the customers in time. There must be means that keep the system operational despite the faults of the instrumentation. Along with the development of the process automation at the control level there is a need for integration of a plant and an office automation systems more tightly. To support business-related decision-making up-to-date, online information about factory floor activities is needed at all levels of the manufacturing enterprise (Pirttioja, 2002). The more information infrastructure of the industrial companies evolves the more visible the trend of the motion becomes from information to knowledge. Industry sees its future benefits in enabling the Knowledge Management solutions. Thus a necessity of development of tools that would facilitate mining, accumulation and sharing of expert knowledge within the whole enterprise is emerging. Such approach can be found e.g. in development course of Metso Corporation (www.metso.com), which is a global supplier of industrial process machinery and systems (Pyötsiä, 2002). The company raises the necessity of Networked Business Environment development through which enterprises all over the world will be able to cooperate. Such environment must facilitate automation of shared business processes (automated partner search, business transactions management, etc.) to make the enterprise fast in decision-making and nimble in direction change. An accent is made on a knowledge-related business when enterprise proposes all the experience that it has gained during its life to an external market. To be available for utilization and reuse, such experience should be kept in a machine (software) understandable form. Realization of these phases deals with a lot of problems, such as compatibility of enterprise information technology standards, fast search of appropriate experiences (services) with quality control, creation of trust enterprise networks (partnership), etc. Development of information support infrastructure for product delivery (logistics challenge) is also important. This direction suffers due to a lack of tools for management of vast material- and information flows from multiple sources to a remote location (Kiianlinna, 2002).

Tools that provide easy monitoring and control of technological processes must be developed. ICTs are actively used in development of such tools, e.g. Plant Automation Systems are developed to provide all necessary information about instrumentation state and control functions. Such systems are the integration of specialized hardware and software components connected in common LAN. Online instrumentation monitoring is actively practiced too. It assumes tracking of real time operating conditions of industrial assets (Sarginson, 2002). This trend leads to complication of field devices towards so-called 'smart devices', which have embedded analytical block (software + hardware). The latter can collect the data about the state of correspondent instrumentation and provide access to it from mobile or stationary terminals (Pyötsiä and Cederlöf, 2000). Usually smart devices can notice their fault states and notify responsible personnel via mobile phone, email, etc. It isn't always economically efficient to equip all field devices with such capabilities and thus only crucial field devices can be enhanced in this way (Sarginson, 2002).

Information infrastructure also evolves to provide efficient communication between components of a Plant Automation System. Digital Field Bus has substituted the analog control systems and there is LAN that spreads across the whole plant (Automation, 2002b). Provision of interoperability between condition-monitoring systems and heterogeneous field devices is an important challenge. This allows users to choose the best of the class products and not to be tied to one manufacturer by a proprietary software system. Although the current level of plant automation is high, there is a lack of intelligent features in it. Industry needs intelligent tools able to derive necessary knowledge from information about histories of field device conditions that is gathered during a long term through Field Bus. Unfortunately decision-making based on collected data is still the task for human and the workload of operators and maintenance people increases (Pyötsiä and Cederlöf, 2000). Therefore, the challenge is to create decision support systems that would partially or fully substitute a human.

During the past ten years, researchers have been applying agent technology to manufacturing enterprise integration and supply chain management, manufacturing planning, scheduling and execution control, materials handling, inventory management and developing new types of manufacturing systems such as holonic manufacturing systems (Shen et. al., 1997). Recent progress in electronics and ICT has made possible to supply every sensor, actuator, or network node with an extended processing power that provides an excellent support for implementation of distributed embedded intelligence (Pirttioja, 2002). It creates an excellent basis for development of the Field Agent (or "smart device") concept, which assumes equipping field devices with software and hardware components that provide the executing platform for Software Agents. This approach allows combining hardware (fast execution) and agent (flexibility, intelligent functioning, and high-level communication) advantages into the field device implementation. Agent-based systems that perform intelligent condition monitoring and diagnostics within a plant were described in (Mattila, 2001). It was claimed that efficient condition monitoring and diagnostics could be performed by combination of different methodologies, each implemented by a separate agent. Such approach leads to development of multi-agent system where a comprehensive task (after its decomposition) is solved by distribution among members of agent community.

XML-based standards have been actively applied in industry becoming a good basis for Agent Technologies (Siltanen et. al., 2001). XML offered a fairly simple technical solution to the automated process of document combination. Software agents can reuse documentation in such format further for different purposes. To enable automated knowledge management the approaches that utilize machine understandable information must be applied. Nowadays an emerging Semantic Web technology seems to have enough potential to address the discussed challenges. Semantic Web assumes semantic annotation of network resources to make them understandable for applications and agents. Semantic Web successfully facilitates interoperability between intelligent tools. It uses a concept of Ontology to provide mutual understanding between applications on the level of notions.

On the other hand, the emerging peer-to-peer solutions particularly suit fine the increasingly decentralized nature of present-day organizations, whether it is a single enterprise or a dynamic network of organizations. They make it possible for different participants (organizations, individuals, or departments within an organization) to maintain different views of the world while exchanging information. At the same time, because they rely on a keyword search and rather simple knowledge representation techniques, today's peer-to-peer solutions are extremely limited. They cannot easily support the introduction of the new concepts, make it difficult to determine whether two terms are equivalent, and generally can only support very limited levels of automation. Semantic Web technologies have been shown the potential to support all these missing types of functionality (SWWS, 2002).

Thus the combination of concepts provided by Semantic Web and Peer-to-Peer together with modern agent technology seems to be a good basis for the future of industrial application integration, knowledge management, automation and control.

In Section 2 the phenomena of enterprises integration and future visions in industry are considered, which relate to emergency of Intelligent Field Agents and Web Services for such type of agents. The importance of Semantic Web and ontologies in provision of interoperability between enterprises is mentioned. In Section 3 a brief introduction to P2P communication model is given. Some features and advantages of P2P solutions are described. The benefits of joint application of P2P and Semantic Web in integration of Web Services are analyzed. Finally the problems concerning P2P + Semantic Web combination are listed. In Section 4 the case of integration and interoperability provision in decentralised industrial collaborative environment of the OntoServ.*Net* project is described. In Conclusions the summary and future research plans are presented.

2. SEMANTIC WEB: PROVIDING INTEROPERABILITY AND INTEGRATION

Very few business applications can live in isolation. Most often, applications have to be integrated with other applications inside and outside the enterprise. This integration is usually achieved through the use of some form of "middleware". Middleware provides the "plumbing" such as data transport, data transformation, routing etc. (Basex, 2002). Enterprise integration can also help to reduce costs, increase operational efficiencies, expedite time to market, and improve return on information technology investments. Without enterprise integration, various infrastructures will lack the robustness and flexibility in a dynamic economy (Hohpe, 2003). Now integration is the direction of success for enterprises. They invest large portions of their ICT budget on informational integration.

Recent Web developments allow making good business by launching Web services, which are connection points through which a client can buy experience that a company has accumulated. Thus the problem of enterprise information infrastructure integration results to problem of integration of their Web Services.

Consider also communication between industrial information agents and industrial Web Services. Obviously to understand each other they must use the same vocabulary of terms. Ontological approach tries to address this problem. Ontology proposes a way of formal description of problem domain to enable automated processes where understanding of content is needed. By standardizing terminology, industrial ontologies enable interoperability between distributed intelligent systems. However, if different industry enterprises separately engineer ontologies for the same problem domain, then we again have interoperability problem, but this time at an ontological level. A mechanism of inter-ontological relations establishment should be developed to enable such interoperability.

Industrial enterprises realized that their successful integration can’t be performed without providing interoperability of information systems being in use. Some challenges related to the emerging Semantic Web applications have been discussed at the Silicon Valley World Internet Center Symposium held on December 11, 2001 (SVWIC, 2002). Since Semantic Web provides interoperability via common inter-enterprise taxonomies (ontologies), a challenge would be to join efforts of enterprises for development of them. As it was explained above, the ontologies represent the entire set of relationships within an enterprise (Benjamins et. al., 2002). The development of ontologies must be performed in a coordinated manner to ensure the consistency of the whole model. Taxonomies are defined by the participants as the vocabulary used inside of the ontology. Language and terms need to be common among users for the data to pass through the various relationships of the ontology and to mean the same thing to all parties. Successful applications of ontologies in decentralized and distributed environments require substantial support for change management in ontologies and ontology evolution.

3. PEER-TO-PEER

3.1. P2P communication model

In (SWWS, 2002) peer-to-peer (P2P) is defined as a class of applications that takes advantage of resources –storage, cycles, content, human presence- available at the edges of the Internet. The peers connected to the peer network together make up a system as whole. A peer could be a computer, a personal mobile terminal or some other device. In P2P systems, computers (that earlier were just acting as clients) now act also as servers. Which role a resource plays from one moment to the next one, depends on what are the systems needs. This takes away the heavy load and dependency of some of the individual servers (Åkerström and Gröndahl, 2002).

If we look to the existing systems on the Web, we can roughly divide them in three different models: the broker mediated model, the direct P2P model and the resource-sharing model.

In a *mediated* P2P network, central servers contain an index of all the content and where it can be found. When the server receives a request, peers hosting the content are identified from the index.

Direct P2P model let users register information with the network neighbors. Information search across the network is based on sending queries to neighbors, and if the neighbors do not know the answer, they send the query to their neighbors or a subset of neighbors. Techniques, like a history profile of the query, could prevent cyclic behavior and restricts the chain length.

In a *resource-sharing* model, a "master" uses "slaves" for any kind of purpose. The master could, for example, use the idle CPU cycles from the slave for calculating extraterrestrial data (SETI@home, Popular Power), using disk space (Mojo Nation) or using the data stored on the slave (like Google: it visits sites to get information and include it in its database).

The P2P model could be seen as an alternative to the client-server model that is mostly in use in today's networks. Since client-server solutions have been used for a long time they are studied better and standardized. The centralization also makes easier configuration and control of performance, security and reliability. Client-server systems, on the other hand, have limitations concerning scalability and are often very costly to own (Åkerström and Gröndahl, 2002).

3.2. Peer-to-Peer: enabling Distributed Content Management

There are three key trends driving the need for, and emergence of, distributed content management solutions: explosion of unstructured data; the critical need to formally manage content; and internetworking and collaboration within and between enterprises. These trends are converging to produce two key requirements—the need to create superior online user experiences and the need to work collaboratively.

Distributed content management systems address the need to access content wherever it resides, produce content while maintaining control over it, and collaborate efficiently by sharing data real-time within a distributed network of stakeholders. These systems create virtual content repositories that eliminate the need for structured storage. In fact, with these systems, data structure becomes irrelevant because information is accessed at its source, in its native format, expanding the reach and participation of stakeholders.

Gartner Research Group (GartnerGroup, 2001) sees immediate synergies between distributed content management and Web content management (WCM) solutions. Together, distributed content management and WCM solutions provide access to potentially all enterprise and interenterprise content, and allow content to be effectively managed and distributed via the Web. Given e-business trends, market drivers and the importance of partnership strategies, Gartner expects that P2P content networks will become prevalent in future. Gartner also believes that half of the current server-based content management vendors will add Data Centered P2P functionality to their product offerings by 2005 (0.7 probability).

3.3. P2P and Semantic Web – emerging Knowledge Management approach

In today's knowledge-based economy, the competitiveness of enterprises and the quality of work life are directly tied to the ability to effectively create and share knowledge both within and across organizations.

The combination of Semantic Web and Peer-to-Peer is highly innovative with prospective benefits to the individualization of work views as well as to the facilitation of knowledge sharing. We want to discuss in more details, why the integration of these two approaches really provides some new opportunities.

Peer-to-Peer computing combined with Semantic Web technology will be an interesting path to switch from the more centralized KM solutions that are currently implied by ontology-based solutions to a decentralized approach. P2P scenarios open up the way to derive consensual conceptualizations among employees within an enterprise in a bottom- up manner (SWWS, 2002).

While in the server/client-based environment of the World Wide Web metadata are useful and important, for the P2P environments metadata is absolutely crucial. Information Resources in P2P networks are no longer organized in hypertext like structures, which can be navigated, but they are stored on numerous peers waiting to be queried if one knows what he/she wants to retrieve and which peer is able to provide that information. Querying peers requires metadata describing the resources managed by these peers, which is easy to provide for specialized cases, but non-trivial for general applications (Neidl et. al., 2003).

Ontologies have shown to be the right answer to knowledge structuring and modeling by providing a formal conceptualization of a particular domain that is shared by a group of people in an organization. However, Knowledge Management Systems, which are based on centralized ontologies, need a long development phase and are difficult to maintain. From a technological point of view P2P solutions are particularly well suited, because they make it possible for different participants (organizations, individuals, or departments) to maintain their own knowledge structure while exchanging information (SWAP, 2003).

Semantic Web and Peer-to-Peer (SWAP) project in (SWAP, 2003) demonstrates that combination of these successful technologies will allow support for decentralized environments. Participants can maintain individual knowledge structures on their peers (PCs), while sharing knowledge in ways such that administration efforts are low, but knowledge sharing and search is easy.

The system developed in SWAP project consists of a set of peers called "SWAP Nodes" (see Figure 1). The knowledge of a particular peer is extracted from several Knowledge Sources, then integrated and stored in the Local Node Repository (LR). A user interface ensures that a user is able to edit/browse/query the knowledge. Queries that cannot be answered by the available knowledge are sent to the whole system. A specialized component deals with rewriting these queries and selecting the peers, which are likely to know the answer.

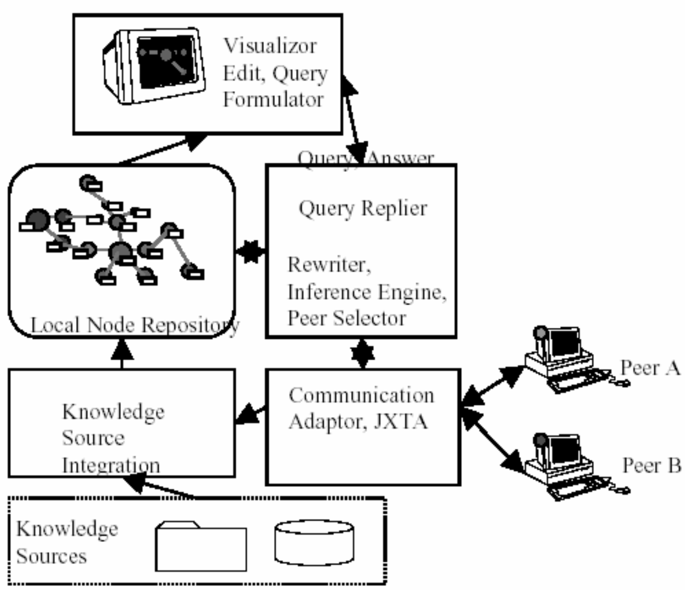

Figure 1. SWAP system architecture (SWAP, 2003)

Every participating entity in the SWAP system has to be enabled to provide its knowledge. In appropriate module, the user can make selection among existing knowledge sources from his/her personal computer for addition (emails, files, folders, bookmarks, or whole ontologies). A special component will extract ontology-like structures from the selected sources. These sources will then be integrated automatically into the LR, which is represented in RDF(S). Merging routines will be required at this point. Content is therefore stored as an ontology, which allows high-quality processing. The LR is an integrated view on all known information sources, both local sources and other peers. Any change in the knowledge sources is propagated to the LR. Queries can be entered by clicking in the views graph or manually as text. If the inference engine cannot get an answer from the local repository, it splits the query and distributes the subqueries in the P2P network.

3.4. P2P and Semantic Web – integrating Web Services

Today distributed computing is connected with the idea of Web Services. These are (parts of) programs that can be accessed over a network using well-defined protocols. An important aspect is that the interactions should be done automatically by computers. Currently one of the main problems is the locating of Web Services, which provide the desired functionality (Thaden et. al., 2003).

Universal Description, Discovery and Integration (UDDI) is a specification for business registries from Ariba, IBM and Microsoft. It is a central (possibly replicated) registry, which contains information about businesses and their provided Web Services (see Figure 2).

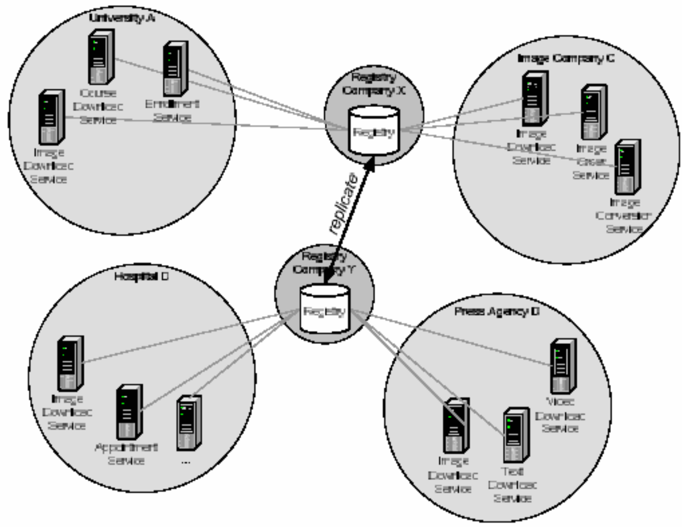

Figure 2. Services registered at central registries (Thaden et. al., 2003)

Centralized approaches combined with the replication have many drawbacks, e.g. poor scalability and less consistency on large registries. UDDI can be seen as typical yellow pages, pointing to registered Web Services, which can be located elsewhere. Furthermore, the search facility is limited; one can search by keywords, but cannot ask for “something similar” since UDDI does not provide a vocabulary (ontology). This situation makes it difficult to discover services. Technically, it would be possible to replicate registry information from all private registries to the public central nodes. However, this needs a replication contract between both registry providers, and manual system administration for each new private registry. Therefore, while technically possible, practically replication from private to public UDDI registries doesn’t occur.

To alleviate this problem, researchers in (Thaden et. al., 2003) after careful analysis of related work done by other groups proposed solution based on combination of Semantic Web with P2P. They suggest moving from a central design to a (de facto already existing) distributed approach by connecting private registries with P2P technology. The power of the Internet comes topmost from moving from centralized solutions to non-structured information based at each computer connected to the net. P2P architecture follows this idea and adds value by creating a virtual global registry from all connected local registries. P2P enables the Web Service registries. Thus, companies as well as universities can build their own Web Service registries, which are maintained by themselves. Being a peer in a P2P-network makes it easy to search all local registries.

A similar project to be mentioned is Speed-R, which is led by University of Georgia (Sivashanmugan et. al., 2002). As (Thaden et. al., 2003) states, in (Sivashanmugan et. al., 2002) researchers have developed a distributed registry based on UDDI without semantically enriched Web Service description. While (Thaden et. al., 2003) envisions a registry partition where institutions maintain their private registries, the Speed-R system assumes a partition based on business domains. The most important difference is that (Sivashanmugan et. al., 2002) uses DAML-S as service description language instead of UDDI tModels. This is the prerequisite for enhanced semantic search capabilities.

3.5. P2P: enabling communication between industrial Field Agents

While Semantic Web provides interoperability between Industrial Agents, data and knowledge sharing between industry enterprises (or plants within enterprise) is provided by communication between agents. This approach is very effective and natural. Agents can be either wrappers or emerging Field Agents.

Use of common ontology allows knowledge exchange between agents, expressed in unified terminology. Knowledge that agent can share with others can be represented in machine-understandable form based on RDF.

Agents can travel from node to node (mobile agents) if a close communication is required between nodes. Mobile agents have many characteristics that enable them to enhance managing control and alarms in territory-wide systems. Mobility is obviously one of the most important capabilities and recent solutions already utilize it. For instance in (Quintero et. al., 2002), a multi-agent approach was successfully applied for building knowledge-based system for efficient management and monitoring of urban infrastructures. They utilized mobility to enhance control and alarm management in urban infrastructure.

Agents also can get the problem from the users or other agents, discover needed resources, consult with other agents (negotiation) and offer a proper solution. They also learn from the past, update their knowledge and predict the future events (ACM, 2001). The main difference between agents and ordinary software is the issue of coordination, cooperation and learning. Agents work together, use the distributed resources optimally and work as a team to solve problems. Agents are flexible entities and are capable to adapt themselves to new environments. That means that utilization of the agent technologies is very suitable for dynamic platforms with distributed components. Although agents are dependent entities, they always communicate with other agents to discover new resources they need. While designing a system based on agent technology, all the characteristics of the autonomous and intelligent agents have to be considered to take most advantages of them (Homayounfar et. al., 2002).

If integration of enterprises is made on the basis of some trade agreements, elements of commerce will certainly be present in the information communication. For example, assume that one have to pay in order to get some knowledge from some agent. Currently, commerce is almost entirely driven by human interactions; humans decide when to buy goods, how much they are willing to pay, and so on. But in principle, there is no reason why commerce cannot be automated (Jennings and Wooldridge, 1998). As an example, Chavez and Maes (Chavez et. al., 1996) describe a simple 'electronic marketplace' called Kasbah. This system realizes the marketplace by creating 'buying' and 'selling' agents for each good to be purchased or sold respectively. Commercial transactions take place by the interactions of these agents.

The multi-agent approach can be successfully applied to the Business Process Management. Company managers make informal decisions based on a combination of judgment and information from many departments. Ideally, all relevant information should be brought together before judgment is exercised. However obtaining pertinent, consistent and up-to-date information across a large company is a complex and time-consuming process. For this reason, organizations have sought to develop a number of IT systems to assist with various aspects of the management of their business processes. Project ADEPT (Jennings et. al., 1996) tackles this problem by viewing a business process as a community of negotiating, service providing agents. Each agent represents a distinct role or department in the enterprise and is capable of providing one or more services. For example, a design department may provide the service of designing a telecom network, a legal department may offer the service of checking that the design is legal, and the marketing department may provide the service of costing the design. Agents, which require a service from another agent, enter into a negotiation for that service to obtain a mutually acceptable price, time, and degree of quality. Successful negotiations result in binding agreements between agents. This agent-based approach offers a number of advantages over more typical workflow solutions to this problem (Jennings and Wooldridge, 1998).

Well, finally we have a decentralized network of industrial agents that resulted from spontaneous enterprises integration. Now the concern is which communication model can be appropriate to enable effective communication processes between the network nodes? Such model must be launched in short terms, without radical modifications of network nodes and without building complicated communication management systems.

Researchers have studied these questions and some convincingly claim that P2P architecture can be an environment in which abilities of agents are fully utilized. Agents in their turn are able to improve functionality of a P2P system (Homayounfar et. al., 2002). They also soundly claim that agent technology is the crossing point where AI and distributed systems meet each other. The possible solution to the current drawbacks of the P2P approach is to use agent technology. Autonomous Agents are capable of performing an advanced (dynamic) P2P networking in which nodes (agents) behave intelligently (negotiate, learn, predict, cooperate and etc) somehow that P2P functions may be optimized (Homayounfar et. al., 2002).

P2P model combined with agent approach is an environment in which agents interact and negotiate with each other directly. In (Homayounfar et. al., 2002) this was named as Agent-to-Agent (A2A) communication model, where each node of the network can be a host for one or more agents. Each agent can have a point-to-point communication with other agents within the network. An A2A system is an advanced version of the Internet agent technology, in which agents have more flexibility and efficiency compared with the ordinary agents' models. In fact, A2A architecture is an agent-based model designed and implemented by advanced P2P features. An autonomous A2A design of a P2P system may overcome the limitations of the current P2P applications and improve the efficiency of the components.

In the described vision of future architecture of industrial agent community each member is self-interested. This community therefore develops itself autonomously and do not need any administration. Similar model is a basis of our human world and makes it evolve permanently. The features of human society could be fully adopted in agent communities as soon as their members become intelligent enough.

4. INTEGRATION AND INTEROPERABILITY IN A DECENTRALIZED COLLABORATIVE INDUSTRIAL ENVIRONMENT OF ONTOSERV.NET

4.1. Decentralized Knowledge Management in a collaborative industrial environment

The above analysis of the Semantic Web technology combined with the Peer-to-Peer communication model in the context of industrial enterprises' integration allows to clarify the state of the affairs in this domain and to form a vision for Global Knowledge Management across industrial enterprises. To enable automated knowledge management in collaborative industrial environment all the resources must be semantically annotated. Automated knowledge management concerns 'passive' resources (electronic documentation, media, and hypertext) and 'active' ones (web-services, agents, software). To provide this possibility, the appropriate engineers must create correspondent ontologies, where the relations between all types of resources must be defined explicitly. The ontologies will describe possible collaboration scenarios between the active resources. The passive resources, being enriched by semantic data associated with such ontologies, will enable intelligent operations to be performed upon them by active resources.

It would be extremely hard to design software for such collaborative environment. Client/server centralized administration tools will fail because of their complexity and extremely high demands for computing resources. The above analysis however shows that a multi-agent system can be better appropriate for modelling such global environment. Agent approach allows separating the environment and its inhabitants (agents). These two are independent, and agents can be easily placed in such environment and removed. This approach makes systems highly modular. Legacy software and web-services can be wrapped by agent "shells", which contain necessary components according to an agent paradigm (receptors, effectors and ontology-based inference engine). New active resources can be implemented as pure agents and placed in this environment. Semantic data is simply attached to the existing resources. The administrative efforts for this environment would be minimal. The intelligent agents can be programmed to target some goal, with no need to specify how to achieve it. Their behaviour is quite flexible and the environment becomes self-organized with some amount of uncertainty. Since the multi-agent system is decentralized, the P2P concept fits very well with the agent communication model. P2P model is highly scalable, fault-tolerable, and self-organizing. Taking into account that the communication between agents is P2P-like, i.e. naturally decentralized, it would be reasonable to create central ontology server(s) for providing interoperability. However there is no doubt that there will be diversity of ontologies in such Global Collaborative Environment and it will be very difficult to coordinate standardization in it. Shared ontologies will provide interoperability between information systems within the community of enterprises that have partner agreement. Recently every partner community develop ontologies for its specific needs and do not care about other ontologies. So, in this case client/server model can be applied since the number of members in every community includes just small part of all integrated participants. The shared ontology provides common terms vocabulary and problem domain representation. Once two communities have agreed to collaborate they must establish some mapping between their ontologies.

Our "Industrial Ontologies" (IOG, 2004) research group carries out a case study of Global Network of Maintenance Web Services for Smart Devices ('OntoServ.*Net*'). Detailed simulation of certain scenarios in it allows disclosing possible problems concerning development of such systems. In our case we consider Web-services for *smart industrial devices* that can be used by *field agents* (agents that inhabit embedded platforms of smart devices) for classification of instrumentation faults. The general idea assumes making experience accumulated by every industrial plant (even device) during its life available for whole community of enterprises (devices). Experience is accumulated and stored in a semantically rich format to be consumable by agents. Maintenance Centres are connected into World-Wide Network and a representative of certain Maintenance Web Service (agent) comes to the site of the client when appropriate. The mobile services are needed if the communication with client-side platform is required to be fast, intensive and secure (Kaikova et. al, 2004).

4.2. Ontology-based integration of industrial applications

The research made within the OntoServ.*Net* case study has lead so far to the following results. On a general level all industrial maintenance applications are interpreted as Service Providers (SPs). The possible services are: provision of information about current state of some field device (values of parameters); alarm system (determines degree of danger for current state of field device); diagnostic system (makes a decision about the class of a situation to be occurred). The current values of field device parameters are usually stored in the Real-Time Database (RTDB) that belongs to SCADA (Daneels and Salter, 1999) system. Such services after integration of industrial enterprises become shared and the challenge is to make possible the utilization of one service by any Service Consumer (SC) in the integrated network. SC can be other industrial application wrapped by software agent, which supplements it by communication, reasoning and other agent functionalities.

It seems that the only way to provide the compatibility of input/outputs of heterogeneous SCs and SPs is an *OntoAdapter*. The latter implies some software-dependent piece of code, which is integrated with an industrial application. On the one hand it can perform necessary function calls and data retrieval and on the other hand it standardises the interface with the application according to the Semantic Web concepts. That is, OntoAdapter converts data and function calls from RDF format to application-specific format based on certain industrial ontology and vice versa (see Figure 3).

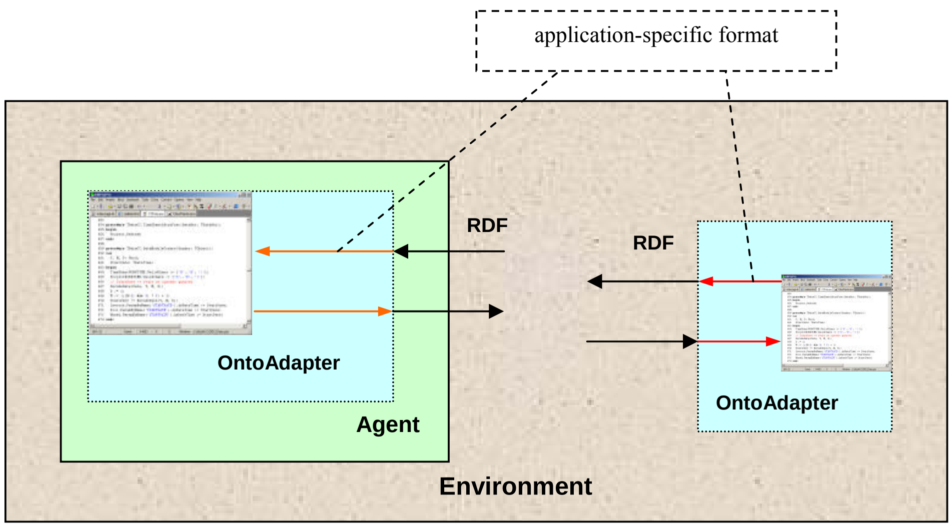

Figure 3. Provision of interoperability between industrial applications

Given the concepts described above, the interaction between SC and SP within single peer will look like the following. First of all assume that in our case the interoperability between industrial applications is provided by some ontology. A simple example of such ontology is given in Figure 4. Three ontologies are described in RDFS files: *ServiceConsumers.rdfs*, *ServiceProviders.rdfs* and *DataTypes.rdfs*. The notions of each ontology are bounded by dotted rectangles in Figure 4 and we also can see inter-ontological relations (relations *needs, provides, in, out*). In the Figure 4, **S** denotes 'subclass-Of' relation. We can see that ontologies *ServiceConsumers* and *ServiceProviders* are independent but both have relations with *DataTypes* ontology. This allows developing ontologies for SPs and SCs independently. But still these two ontologies must conform to standards of data types defined in *DataTypes* ontology. So integration of SPs and SCs is performed via common data types and the system (environment), which possesses the three mentioned ontologies, can make this integration.

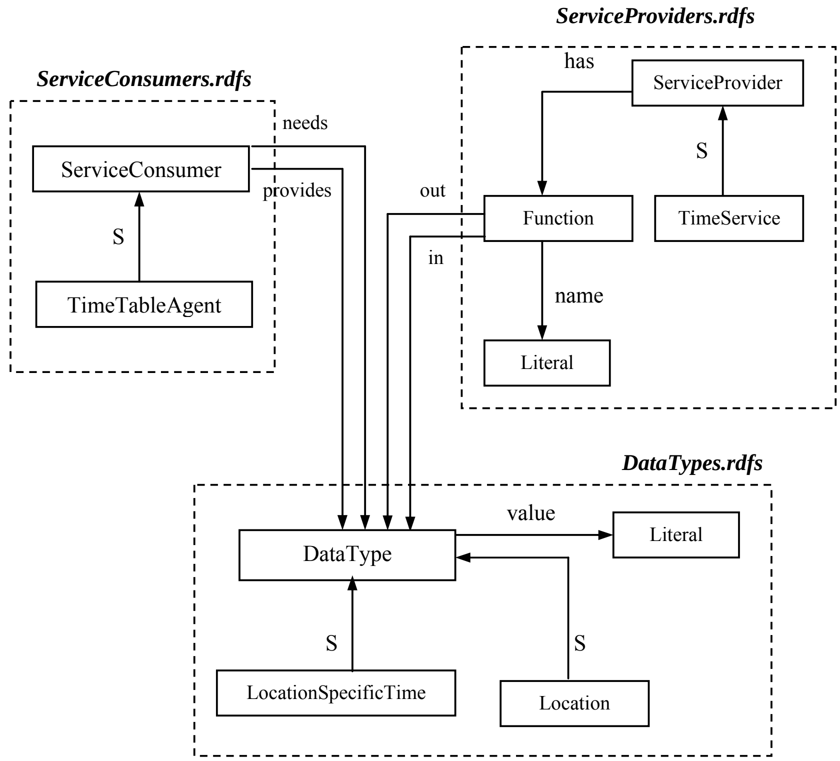

Figure 4. Sample of ontology integration

Let's consider some scenario of such integration. Assume that two entities are registered in the repository of the Environment.

The first entity is *TimeTableAgent* – a software agent, which schedules some worldwide activities for a user (see Figure 5).

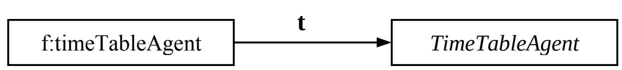

Figure 5. *TimeTableAgent* description

The second entity is *TimeService* – a web service, which allows knowing the current time in the given geographical location. Thus the Environment has a description of *TimeService* in the following form (see Figure 6). In these figures, **t** denotes 'type-Of' relation, **n** is some namespace that uniquely identifies resources. Symbol **f** in Figure 5 also denotes some unique namespace.

Now, assume that for some reason our *TimeTableAgent* needs to know the current time in Ghana. He sends request to Environment in RQL-like form (Karvounarakis et. al, 2002) (see Figure 7).

The request means: "I need a value for the variable *LocationSpecificTime* data type". The question mark denotes omitted data. The environment searches in the repository of entities a web service, which would provide such value. From the three ontologies that were mentioned above the Environment knows that it must find such service, whose *function* has exactly the same *out* as it was requested. So, the search query generated by Environment will be the following (Figure 8) also in RQL-like form. The query means: "*What is the resource identificator of an entity that has type ServiceProvider and has function with an output of LocationSpecificTime type*?". As a response the Environment must get from RQL-like engine (Broekstra and Harmelen, 2002) n:timeService – resource identifier of the *TimeService* and also types of inputs that service requestor (*TimeTableAgent* in our case) must provide. The required function has an input of data type *Location*. So, Environment will send to *TimeTableAgent* the request, shown in Figure 9. The request to *TimeTableAgent* can be translated like: "*Can you provide value for data having type Location*?".

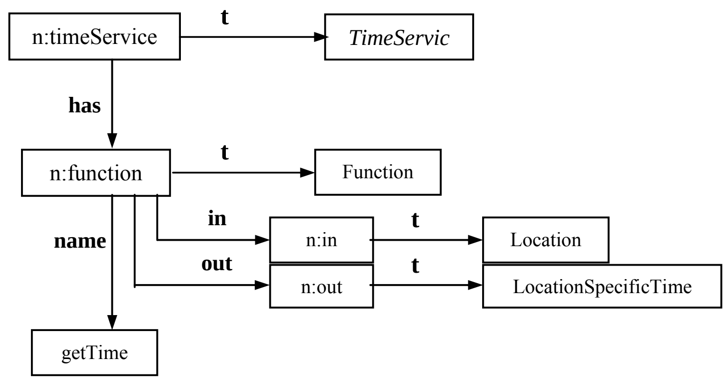

Figure 6. *TimeService* description

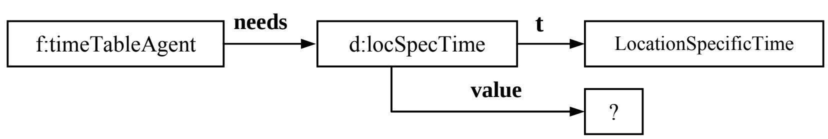

Figure 7. Query example

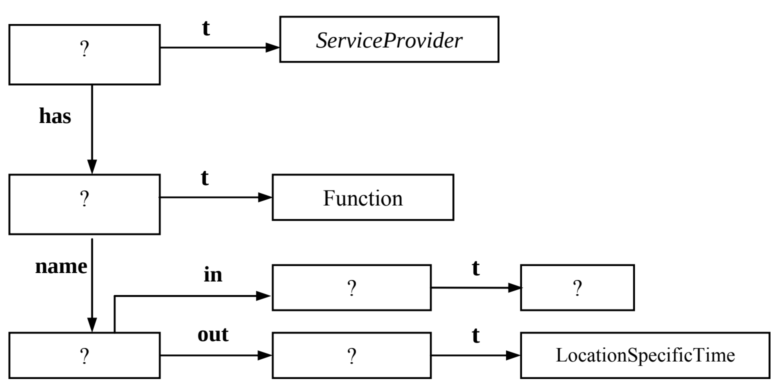

Figure 8. Search query for web service

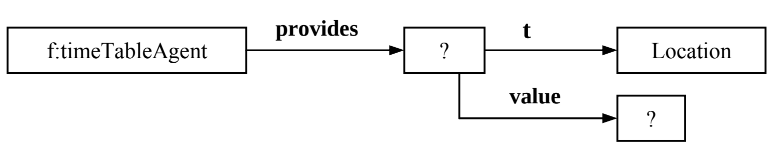

Figure 9. Query to *ServiceConsumer*

If the *TimeTableAgent* provides the value of such type, Environment will send an RQL-like request to *TimeService* to obtain the current time in Ghana (see Figure 10).

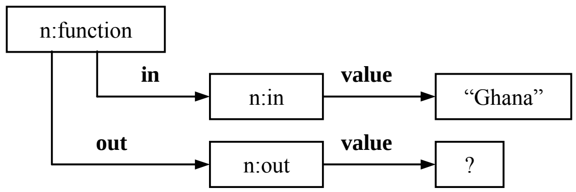

Figure 10. Web service function call

4.3. OntoShell in Peer-to-Peer network

In the previous subsection a scenario of industrial applications integration was described within a single network node. Analysing interaction of many nodes interconnected into decentralized network (peer-to-peer) a concept of *OntoShell* has been developed within OntoServ.*Net* project. According to this concept, every peer in P2P network is represented by correspondent OntoShell, which generalizes and hides its internal structure. Such peer OntoShells are interconnected with each other as neighbours forming P2P network (see Figure 11).

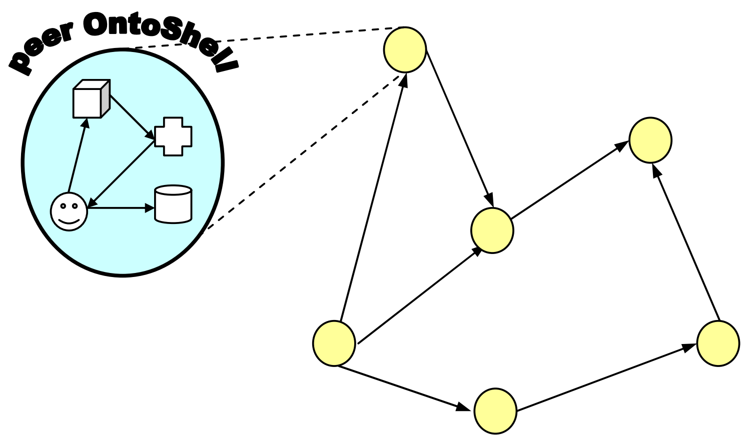

Figure 11. OntoShell generalizes internal structure of a peer

If to assume that each OntoShell accumulates knowledge about optimal query routes (routing information is semantically enriched) one day a group of peers can make a decision about rearrangement. Such rearrangement means formation of peers’ cluster: peers are joined in a group according to some common features. For instance, peers, which provide different maintenance services for Control Valves can form a cluster ‘Control Valve Services’. And this cluster generalizes the features of its members generating meta-profile for a group as for single entity. Inside the group the nodes can be rearranged into centralized topology for more efficient query routing. The duties of a central node can be handed to some peer. And this peer will be an entry point to this cluster; it will possess the functionality of the OntoShell. Thus, OntoShell will be a virtual conjunction of a set of peers (see Figure 12).

After formation of OntoShell cluster, the central peer can still have links with its former neighbours and even can be a member of another cluster. Such clustering of peers will reduce first of all the unnecessary roams of queries. If the query doesn’t match the profile of a cluster it wouldn’t match any profile of its members. So the number of matching processes is decreased by number of members of the cluster.

The challenge here is a process of generation of a cluster meta-profile from the profiles of its members.

The concept of an OntoShell in this context can be also used for integration of formerly independent industrial intranets. Every intranet has its own unique message protocol. In this case OntoShell is “put” on each Intranet and every OntoShell translates messages from internal format into common format for all OntoShells. In this case a single node, which implements the functionality of

OntoShell, must be created for every Intranet. This will be an entry point of external messages into Intranet (see Figure 13).

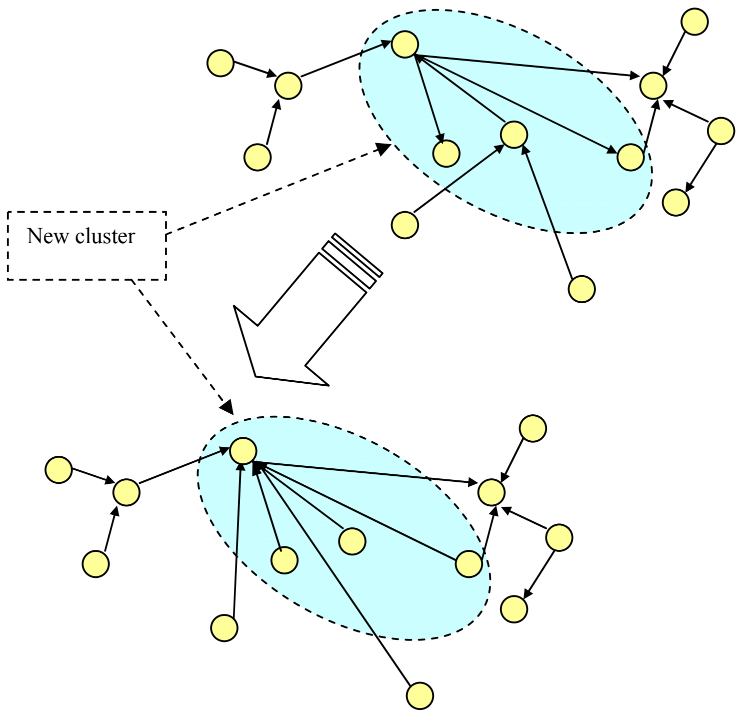

Figure 12. Group of peer OntoShells rearranges into cluster OntoShell

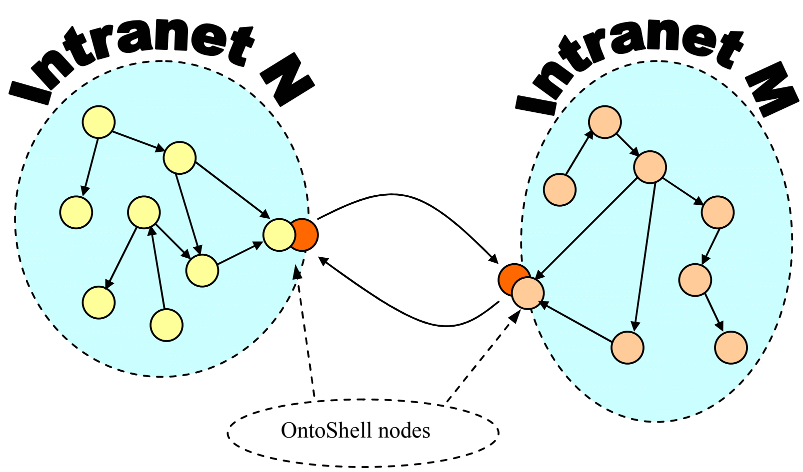

Figure 13. OntoShell provides interoperability between Intranets

5. CONCLUSIONS

To be competitive on a global market modern industry must meet some important challenges. On a plant scale such challenges are concerned with development of highly efficient Automation Systems that would allow controlling manufacturing processes optimally. Survey of systems, which are nowadays exploited on industry plants, allows gaining an understanding of actual state of affairs in this domain. The trend was discovered that the advanced Automation Systems are transforming nowadays into the multi-agent systems. In such systems the community of agents is expected to perform cooperatively condition monitoring, decision-making and process-control tasks. Such systems are self-organizing, intelligent, self-learning and fault-tolerable. It turns that Field Devices evolve into Field Agents and there is some future vision based on this concept. Such a vision assumes that every community of Field Agents (e.g. within a plant scale) would have access to a common Field Agent Server, which accumulates the experience of the agents and makes this experience available to others. This vision leads to the idea of Information Services for Smart Devices. Nowadays one of the most important issues for industry is integration of enterprise information systems and provision of interoperability between them. Here the challenges of effective inter-enterprise Knowledge Sharing and interoperability between automation system components, which come from different vendors, have significant value. This approach can be considered as one potential application of the Semantic Web and Peer-to-Peer synergy.

Enterprise integration problem is closely related to the integration of industrial Web Services. Taking into account the future visions of industrial information infrastructure industrial Web Services have been recognized as services for Smart Device (Field Agents). Therefore, industrial enterprise integration results in automated knowledge sharing between agents as representatives of different enterprises. For clear understanding of these knowledge exchange processes the structure of plant agent community has been studied. Since a question of interoperability between industrial agents on knowledge level has a crucial value, the concept of ontology is discussed in this context. The evidences of beneficial utilization of ontologies in the interoperability issues have been given. The benefits they provide are: common vocabulary of industrial objects definitions, explication of expert knowledge, systematization, standardization and meta-level functionality. Some real-life projects where the potential of ontologies was fully utilized for provision of interoperability between information systems were described.

Implementation of Semantic Web concepts to organize all the industrial information resources is expected to improve the resource sharing. To be suitable for automated search and other automated processing every resource has to be supplied by machine-understandable description on a basis of common ontology. Further analysis has shown that industrial information resources are highly heterogeneous. Semantic Web research has not provided enough evidence about feasibility of its challenges yet. The potential of Semantic Web will be fully utilized only when all the semantics (resource descriptions, expert knowledge, whatever necessary) will become explicitly available.

Emerging Peer-to-Peer concepts seem to be quite appropriate and suitable to the increasingly decentralized nature of modern organizations and their dynamic networks. Obviously, Peer-to-Peer communication model would effectively enable the Web of semantically annotated industrial information resources. The set of attractive features of the Peer-to-Peer model includes decentralization, scalability, self-organization and fault-tolerance. The overview of systems of this kind allows observing a significant progress in this domain. The overview of Peer-to-Peer and Semantic Web solutions for Web Services integration has also been made.

Analysis of the development trends for ICT in industry proves an emergency of the “Smart Device” concept, which assumes the implementation of the embedded platforms for software agents. Future vision of information infrastructure of integrated industrial enterprises will be represented by a global multi-agent system where self-interested agents are motivated to exchange knowledge. Peer-to-Peer also ideally fits provision of communication model for this system. Semantic Web will facilitate the intelligent activities of the agents.

The emerging integration of distributed information systems and a huge amount of various resources requires Peer-to-Peer (decentralized) management. Semantic Web, in its turn, provides significant support for interoperability between heterogeneous Peers and support for intelligent operations with resources in such environment. Application of the Semantic Web combined with Peer-to-Peer for provision of effective information resource sharing between industrial enterprises is promising. However to evaluate finally the utility of industrial interoperability based on Semantic Web and Peer-to-Peer combination more industrial case studies have to be carried out.

ACKNOWLEDGEMENTS

Authors are grateful to Dr. Jouni Pyotsia and his colleagues from Metso Corporation for useful consultations and materials. Also we would like to thank our colleagues from Industrial Ontologies Group (Oleksandr Kononenko and Oleksiy Khriyenko) for useful discussions within the scope of this paper.

REFERENCES

(ACM, 2001)	Association for Computing Machinery, Proc. of the Fifth International Conference on Autonomous Agents. May 28 - June 1, 2001, Montreal Canada. Published in New York.
(Åkerström and Gröndahl, 2002)	Åkerström, H. and Gröndahl, J. (2002) *Collaborative development of ontologies in a Peer-to-Peer environment*. Ms. Thesis, School of Information Technology & Electrical Engineering, Brisbane.
(Automation, 2002a)	Metso Automation's customer magazine, (2002) *Automation,* 3*,* 5-10.
(Automation, 2002b)	Metso Automation's customer magazine, (2002) *Automation,* 1, 20-22.
(Basex, 2002)	Basex Inc. (2002) Executive report: Clerity for Enterprise Knowledge Sharing, www.clerity.com (accessed Apr. 2, 2004).
(Benjamins et. al., 2002)	Benjamins, V. R., Contreras, J., Corcho, O., Gómez-Pérez, A. (2002). "Six Challenges for the Semantic Web". Proc. of ISWC2002, June, 2002.
(Broekstra and Harmelen, 2002)	Broekstra, J., Kampman, A., Harmelen, F., (2002) "Sesame: A Generic Architecture for Storing and Querying RDF and RDF Schema", Proceedings of the 1st Internation Semantic Web Conference, pp. 54-68.
(Chavez et. al., 1996)	Chavez, A., Maes, P., Kasbah, P. (1996) "An agent marketplace for buying and selling goods", Proc. of First International Conference on the Practical Application of Intelligent Agents and Multi-Agent Systems, London, UK.
(Daneels and Salter, 1999)	Daneels, A., Salter, W. (1999) "What is SCADA?", Proceedings of the 7th International Conference on Accelerator and Large Physics Control Systems, Trieste, Italy, p. 339.
(GartnerGroup, 2001]	Gartner Research Group. (2001) The Emergence of Distributed Content Management and Peer-to-Peer Content Networks. Engagement #010022501.
(Hohpe, 2003)	Hohpe, G. (2003) Enterprise Integration Patterns. ThoughtWorks Inc.
(Homayounfar et. al., 2002)	Homayounfar, H., Wang, F., Areibi, S. (2002) "Advanced P2P Architecture Using Autonomous Agents". CAINE, San Diego, California, pp: 115-118.
(IOG, 2004)	Official web site of Industrial Ontologies Group, http://www.cs.jyu.fi/ai/OntoGroup/ (accessed Apr. 2, 2004).
(Jennings and Wooldridge, 1998)	Jennings, N. R. and Wooldridge, M. J. (1998) "Applications of Intelligent Agents", in Jennings, N. R. and Wooldridge, M. J. (eds.) *Agent Technology: Foundations, Applications and Markets*, 3-28.
(Jennings et. al., 1996)	Jennings, N. R., Faratin, P., Johnson, M. J., Norman, T. J., O'Brien, P., Wiegand, M. E. (1996) Agent-based business process management. *Int. Journal of Cooperative Information Systems*, **5**(2, 3), 105130.
(Kaikova et. al, 2004)	Kaikova, H., Khriyenko, O., Kononenko, O., Terziyan, V., Zharko, A. (2004) Proactive Self-Maintained Resources in Semantic Web, *Eastern-European Journal of Enterprise Technologies,* **2**(1), 37-49.
(Karvounarakis et. al, 2002)	Karvounarakis, G., Alexaki, S., Christophides, V., Plexousakis, D., Scholl M. (2002) "RQL: a declarative query language for RDF**",** Proceedings of the 11th International Conference on World Wide Web, Honolulu, Hawaii, USA, 592 – 603.
(Kiianlinna, 2002)	Kiianlinna, P. (2002) Heavy Industry Intelligent Project System (HIIPS). Presentation. Helsinki University of Technology, TAI Research Centre.
(Mattila, 2001)	Mattila, M. (2001) Agent in Condition Monitoring and Diagnostics. Helsinki University of Technology, Postgraduate seminar on automation technology.
(Neidl et. al., 2003)	Neidl, W., Wolf, B., Siberski, W., Qu, C., Decker, S., Sintek, M., Naeve, A., Nilsson, M., Palmer, M. and Risch, T. (2003) EDUTELLA: P2P

Networking for the Semantic Web, *Computer Networks Journal (special issue on Semantic Web).*

(Pirttioja, 2002) Pirttioja, T. (2002) *Agent-Augmented Process Automation System,* Ms. Thesis, Helsinki University of Technology, Department of Electrical and Communications Engineering, Espoo, Finland.

(Pyötsiä, 2002) Pyötsiä, J. (2002) Future Care, ICT Development. Networked IT Solutions. Metso Corporation presentation, www.metso.com (accessed Apr. 2, 2004).

(Pyötsiä and Cederlöf, 2000) Pyötsiä, J. and Cederlöf, H. (2000) "Remote Wireless Presence in Field Device Management",. ISA Proceedings 2000.

(Quintero et. al., 2002) Quintero, A., Konare, D., Pierre, S. (2002) A Knowledge-based Approach for Managing Urban Infrastructures, *Knowledge-Based Systems*, **15**(8), 449-463.

(Sarginson, 2002) Sarginson, M. (2002) Online instrumentation monitoring, impact on safety, uptime and process profitability. Metso Automation, Field Control Systems Committee, Exhibition 2002.

(Shen et. al., 1997) Shen, W., Xue, D., Norrie, D. (1997) "An Agent-Based Manufacturing Enterprise Infrastructure for Distributed Integrated Intelligent Manufacturing Systems". Proc. of the Third International Conference on the Practical Application of Intelligent Agents and Multi-Agents, London, UK, 23-25.

(Siltanen et. al., 2001) Siltanen P., Leikas M., Syrjänen T., Peltonen T. (2001) How to use XML and PDM in document production. *XML Europe 2001*. Berlin.

(Sivashanmugan et. al., 2002) Sivashanmugan, K., Verma, K., Mulye, R., Zhong, Z. (2002) SpeedR: Semantic Peer-to-Peer Environment for Diverse Web Service Registries. Final Project, CSCI 8350: Enterprise Integration.

(SVWIC, 2001) Silicon Valley World Internet Center (2001) Data-to-Cash (D2C): Business Opportunities Emerging from an Ocean of Data. Symposium.

(SWAP, 2003) Ehrig, M., Tempich, C., Staab, S., Harmelen, F., Siebes, R., Sabou, M., Broekstra, J., Stuckenschmidt H. (2003) "SWAP: Ontology-based Knowledge Management with Peer-to-Peer". 4th European Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS'03), London.

(SWWS, 2002) Siebes, R. (2002) Peer-to-Peer solutions in the Semantic Web context: an overview. EU-IST Project IST-2001-34103 SWAP, Vrije Universiteit Amsterdam.

(Thaden et. al, 2003) Thaden, U., Siberski, W., Nejdl, W. (2003) A Semantic Web based Peer-to-Peer Service Registry Network. Technical Report.

ABOUT THE AUTHORS

Vagan Terziyan, 1958, received his Engineer-mathematician degree on Applied Mathematics in 1981 from Kharkov National University of Radioelectronics (KNURE). He became Candidate of Technical Sciences (Dr. Tech. equivalent) in 1985 and Doctor of Technical Sciences in 1993 (Dr. Habil Tech. equivalent) at the Software Engineering Department. He is acting as Professor on Software Engineering since 1994 and as the Head of the Department of Artificial Intelligence and Information Systems since 1997 in KNURE. Area of research interests and teaching includes but not limited by the following: Intelligent Web Applications, Distributed AI, Semantic Web and Web Services, Peer-to-Peer, Knowledge Management, Machine Learning. Currently he works as Associate Professor in MIT Department, University of Jyväskylä and as Senior Researcher at the InBCT TEKES Project in Agora Centre and Head of "Industrial Ontologies" Group.

Andriy Zharko, 1981, graduated from the KNURE (Ukraine) in June 2003 with Diploma of Engineer in Intelligent Decision Support Systems. In December 2003 he successfully finished the Master's program at the University of Jyvaskyla, Finland, Dept. of Mathematical Information Technology, where he obtained a degree of Master of Science in Mobile Computing. He is also a member of the "Industrial Ontologies" research group since January 2003 and a researcher/developer, postgraduate student at the Agora Research Center, University of Jyväskylä. His research is concerned with applications of Peer-to-Peer Semantic Web technologies.