Use Case 1 in Transport – Research Challenges
Managing Knowledge at Trenitalia

KW Partner: University of Trento
IB Member: TSF Trenitalia

1. General Description of Business Use Case

TSF asked University of Trento to design a knowledge management system able to support the activities of a major business unit, UTMR (Unità Tecnologie Materiale Rotabile), which is responsible for the design, maintenance and engineering of rolling stocks which are manufactured by external suppliers. Such unit is composed by heterogeneous and specialized professional communities that need to both locally manage their knowledge and, moreover, to meaningfully exchange knowledge across communities and in particular to solve the semantic issues posed by the need to preserve linguistic heterogeneity while facilitating coordination and collaboration.

UTMR is in charge for providing Trenitalia with rolling stock and to maintain it. The headquarter is in Florence, and the core business deals with both the design of new rolling stock projects, and the acquisition / maintenance of new railway material. UTMR employs 5.000 workers distributed, according to their skills, over ten workshops (called Business Units, each specialized on a different kind of rolling stock). These are geographically distributed in Italy and are in charge of periodical maintenance of Trenitalia's rolling stock and materials.
Briefly, this process is made up of the following phases:

- Definition of a new train or train component's features (as the documentation for the setting up of a call for bids)
- Negotiation phase (as the selection of the right provider according to the required features)
- Order managing and checking (as the process of monitoring the provider's work)
- Running (as train's managing and maintenance)
- Leaving (as train's dismissing and, eventually, selling)

The analysis we run made evident that UTMR can be described as a constellation of heterogeneous and specialized professional communities (such as engineers, testers, maintainers, project managers). Moreover these communities are of different kinds: the first corresponds to the different company's functions: Gestione Commesse GC (Order Managing), Tecnica e Ricerca TR (Research and Technology), Sperimentazione S (Testing), Ingegneria degli Impianti e della Manutenzione IIM (Maintenance and Plant Engineering). These communities are official, that is, their existence is formally recognized and legitimated by the organization. Moreover they are permanent, since they do not have a specific goal to achieve, but rather they continuously provide competence to a particular segment of the core process.
Besides permanent communities, there are other communities which are "crosscutting" the formal organization:

Communities by component (permanent), made up of workers who spontaneously aggregate in order to share a common professional interest (see below for more details about this kind of community), like for example the community of pneumatic experts. At the beginning of the analysis, these communities had no formal recognition nor had some legitimate "space" for their collaboration. Nonetheless, they represent an important source of knowledge and, moreover, of contamination across departments since members belong to different functions.

Communities by project (temporal), made up of workers involved in a specific project, such as ETR500, or Pendolino (which are names of new train projects). At the beginning of the analysis, these communities were formally recognized as project teams in charge for achieving some goal. Nonetheless, they where not seen as sources of new solutions and best practices that could be spread and reused in other projects.
Obviously, each UTMR member belongs at the same time to different communities. Such multi-membership is a crucial aspect in order to bridge knowledge across different communities.

2. Proposed Semantic Web-based Solution

The goal of this case study shows the double faced need of centralization and distributedness pragmatically approached not in terms of a dichotomy, but rather in terms of a process that occurs within and across organizational communities. In this process, according to a distributed approach to KM, organizational members have individual interpretations of the world which express their professional and social identities. These interpretations are manifested as local classifications and sharing policies. Moreover, due to their need to meaningfully exchange knowledge, people engage in participation and semantic coordination processes through which the different interpretations can find some temporary and partial relationship. These interpretations represent the official, and consolidated conceptualizations that each community generates in order to stably support intra-community knowledge transactions. At an intermediate level, the individual and the community level are connected by means of a structuring process whereby members contribute and validate local knowledge to make it institutional. This model, here named the “onion” model since composed by these three layers (participation, structuration, institutionalization. see fig 1), has been implemented in the case study using a mix of existing KM tools. Nonetheless, major challenges emerged that can constitute interesting fields of exploration for the semantic web community.

The Maintenance Technical Procedure writing process

Communities are involved in several UTMR activities. For the present business case, we want to focus on the process of writing a Maintenance Technical Procedure (Norma Tecnica di Manutenzione – NTM) as it clearly represents the “onion” model described above.

A Maintenance Technical Procedure (MTP) is a document that workshops’ people access in order to find all the information about maintenance activities required for a specific train, component or railway material. For their criticality, MTPs must be considered as sources of stable, consolidated, institutional knowledge. On the other hand, the production of this document involves several UTMR workers, belonging to both the specific informal communities “by component” that are involved in the maintenance issue, and the particular functional communities that are formally involved in the core process. All these people, since they view the problem from quite different perspectives, need tools in order to retrieve and store knowledge according to their individual and professional conceptual schemas. Moreover, they need to establish temporary and partial semantic relationships with the others in order to meaningfully share the knowledge required to achieve the common task. Among these two extremes, the institutional one of a complete MTP and the informal one of digging through notes and exchanging tips, there is a structuration process through which unorganized content is organized into an increasingly consistent and unitary knowledge object. This flow is organized around the specific working practice of the community, made up of roles, rules, and artefacts.

From a knowledge management point of view, the MTP writing process involves three different levels.

1. PARTICIPATION LEVEL

Searching for experts and relevant information about MTP

At a first stage, the knowledge work consists in searching different sources (both internal and external to UTMR) all the relevant information that could be used to write the specific MTP they are asked to work on, like similar MTPs, reports on maintenance problems, documents on repairing experiences.

The relevant sources and repositories that UTMR people access during the MTP writing process are:

1. GAD (Gestione Archivio Disegni): it is the repository dedicated to the storage of projects' designs provided by external suppliers when the order ends. It is accessed by TR and IIM people for the writing, checking, validation and reviewing of a new MTP referring to train's components whose designs have been stored in GAD.

2. SICUESE: it is an Access database, dedicated to the storage of documents dealing with maintenance activities, including MTPs. It is accessed by Business Units (Workshops) and IIM people.

3. RFI web site: RFI (Rete Ferroviaria Italiana) is the Trenitalia company who manages the railway infrastructure. In its web site, one can find circulating rules and parameters (those relevant from the infrastructure point of view), which are relevant in writing a MTP.

4. RSMS: it is a SAP system where all the repair and maintenance operations performed on each rolling stock item are recorded during its life cycle. It is supplied by BUs and accessed by IIM and TR people in order to have those running data which are relevant in writing a MTP.

Concerning the relevant information search process, the system should provide retrieval tools able to find documents and information dealing with the theme treated by the specific MTP. This means both traditional tools supporting a full-text keyword search, but also more semantic enabled tools. Among these, tools are needed in order to extract from documents relevant parts and, moreover, able to match different categorization structures, since the different UTMR members and communities use different ways to classify documents that refer to a same subject. Obviously this introduces semantic criticalities, which can be solved by means of semantic matching algorithms.

In order to support a responsible for the writing of an MTP, the system must be able to find relevant experts for each technical topic. In doing so, UTMR members should be categorized according to the conceptual schema that underlies an MTP. For example, users should describe their competences and skills by filling a profile form which can then be searched by the system itself to suggest potential community or writing group members. The system should also allow users to build temporary expert communities, that can collaborate in order to solve some specific problem. Notice that such communities and their knowledge can be used also as targets for specific searches needed by other MTP teams. This is particularly crucial in the case of MTP that treat a completely new component. In fact, neither specific experts, nor previous MTPs can be consulted. The knowledge needed to write this kind of MTP is not explicitly available, but it can be created through the combination of different expertises that deal with different topics.

2. STRUCTURING LEVEL

MTP Writing and Validation

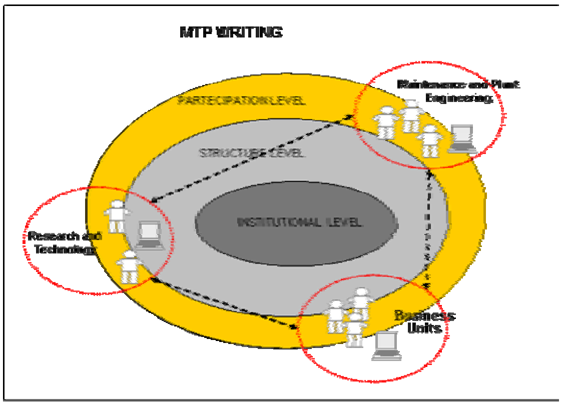
From a logical, but often also a chronological point of view, the collaboration phase is followed by a structuration process. Here, knowledge which is still expressed in a raw collection of partially related knowledge objects, is organized, filtered and validated becoming a consistent and unitary knowledge product. In the MTP writing process, the structure level refers to the phase in which writing groups start defining the structure and contents of the new MTP by using the relevant information previously collected and organized. The system should support this activity by enabling the document workflow and life cycle, and connecting them with functionalities like document versioning management, automatic status changing, check out and check in. Moreover, the system should support the writing of the Activity Report which is complementary to the MTP. A particular aspect of this phase, is that it reflects the social structure and practice of the specific community; that is, each community should be able to define a process and an ontology that reflects its current work practices. In the case of the Activity Report, team members record the entire process they follow in order to get to the new MTP: sources they have explored, people they have contacted, sites they have navigated, other MTP they have read, etc... These activities vary according to the type of MTP, since the process depends on the particular component treated, on the people involved in the team, on the existence of previous MTPs that deal with the same component in other type of trains, etc…All the activities performed by team members are then reported within the Activity Report according to a chronological order. The collection of all the Activity Reports becomes a sort of library of all possible processes people can follow in order to work on a new MTP. In fact the Activity Report collection itself is a quite precious source they can access when starting the formulation of a new MTP. This practice reveals, on the one hand the need of structuring the process and, on the other, the difficulty to standardize work according to a single process definition to be followed each time a new NTP has to be written. As a consequence, a requirement is the one of a workflow engine able to be easily redesigned in order to both match the characteristics expressed by different practices, and evolve with the evolution of a living community. In particular, besides traditional requirements such as workflow management system, document versioning, notification, document tracking, check out and check in, the system should be intended as a “meta-technology” through which users can collaboratively negotiate and design their technology as a solution. Moreover, such negotiation involves not just the definition of the process, but also the definition of a shared conceptual schema. From this perspective, tools are needed in order to support the evolution of local schemas into shared schemas, suggesting to members how their conceptual structure should change in order to “match” the one owned by others.

3. INSTITUTIONAL LEVEL

Publishing the new MTP in the official company repository

The final phase of the MTP writing process occurs when the responsible stores the MTP in the company official repository. At this stage corresponds the moment in which knowledge is institutionalized becoming part of the intellectual capital of UTMR. This means that the system should allow the responsible to upload the final document in a community's categorization structure which has been previously created by the community knowledge manager. The MTP is now available to who ever needs it. Here, besides traditional requirements such as information storing by multiple views, information searching by keyword and by attributes, and navigation of the repository's categorization structures, tools are needed in order to detect when communities' processes definitions or ontologies no longer represent shared agreements. In fact, since institutional knowledge is the expression of a shared language and practice, when changes occur in the broader context, the community engage in a renegotiation phase aimed at reestablishing some agreement on how things should be done. In particular, changes in local conceptual schemas at the individual level which are not compatible with the institutional level, may activate participation and structuration processes through which differences are made explicit, reasons and motivations are explored, and new linguistic and processual configurations defined.

An example of such "changing process" is when people working in the Business Units (the train workshops), due to their practical work, generate new ideas on how to maintain a particular component, or notice an incongruence between the guideline suggested by the official MTPs about the "isolated" component, and the effects produced by its interaction with another component (treated in another MTP) which makes the guideline no longer appropriate. In this case, we observed a sort of "struggle" between on the one side (the "bottom level" - Business Units' workers) the emergence of a changing need and, on the other side, (the "top level" - MTP writers) the resistance to accept the new "element", perceived as a sort of revolutionary change, able to overthrow institutional norms.



The proposed technological architecture

In order to support such process, a particular technological architecture has been proposed. This architecture includes three different applications which are integrated in order to support the three stages of the model described above. It is to be noted that the proposed architecture was not able to match all the requirements, even if some of these where matched through a manipulation of existing solutions that went beyond their original purpose.

1. KEEx: is a Peer-toPeer document sharing system, enabling communication among organizational actors (users, groups, organizational units) while respecting their semantic autonomy and their privacy and security requirements. In particular, KEEx is based on a Distributed Knowledge Management approach whereby (i) each peer (called a K-peer) provides all the services needed to create and organize "local" knowledge from an

individual's or a group's perspective, and (ii) social structures and protocols of meaning coordination are defined to achieve semantic coordination among autonomous peers (e.g., when searching documents from other K-peers). In KEEx, local perspectives are represented by each peer's classification schemas which are used to both organize and retrieve knowledge from other peers. KEEx applications are mostly developed with JAVA. The Peer to Peer infrastructure is based on communication protocols developed within the JXTA project, a SUN sponsored initiative (www.jxta.org).

2. Verity K2 Enterprise: is a quite complex suite of content and information retrieval tools (like, for example taxonomy builder, social network tools, document and content retrieval tools, expert location tools) based on a proprietary search technology. In the proposed architecture it has been included just the lexical search engine

3. FileNet P8: is the FileNet's Enterprise Content management (ECM) platform.

In the proposed architecture the following modules are included:

- FileNet Content Manager. This module, thanks to an "event driven" architecture, allows users to manage the whole document life cycle, from its creation to its storing, from its dissemination to its reviewing, from its publishing to its reuse. Document management services could be installed on a single server or could be distributed over more servers. FileNet Content Manager is based on standard platforms, server and network protocols.

- Filenet Business Process Manager. This module allows users to design and manage document and communication workflows.

We underline that the architecture reflects two sets of major requirements:

1. the needs that a single community has in working on its core activities according to their local practices. In the proposed example, we implemented the collaborative practice (the MTP writing) that characterizes a specific type of community (communities that work "by component"). The project has also designed community support processes for other groups (such as project teams) that are characterized by different working practices (e.g. the designing of a new rolling stock). In this case, KM requirements are similar to the ones of the MTP process, even if the weight of each layer is different. Infact, since the order core process is more structured, the knowledge process focuses more on the second and third level (structuration and institutionalization) rather than on the first (participation). Another configuration is the one that we can find in the third kind of community, the one corresponding with company's functions; here the first level is particularly critical, since community members need to freely collaborate in order to share the different experiences they produce while working on the different projects. In this case the focus is not on producing something fix and official, rather is on knowledge sharing in order to improve the whole community's competence. We can say that according to the type of community, the extension of each level may vary. These different needs are represented in the proposed architecture which is modular and scalable.

2. the needs of coordination expressed by the organization when viewed as a constellation of heterogeneous and specialized communities. These communities need autonomy in order to consolidate their competence. On the other hand, meaningful knowledge exchange processes are needed in order to foster coordination and innovation. In fact one of the main problems faced bu UTMR was the weak collaboration dynamics across professional

communities, in particular the formal ones, that used to work autonomously even when “forced” to collaborate on projects that where typically multifunctional. This architecture supports the emergence of collaboration dynamics thanks to the idea of supporting multimembership, the emergence of cross cutting communities, and semantic interoperability across heterogeneous classifications by means of semantic matching enabled tools (KEEx).

Open Points

As said above, a major goal of the proposed KM system was to support the needs that communities express to work on their core activities according to their local practices. In particular, this meant the capacity to provide a solution able not only to represent each practice once, but rather to continuously evolve with practice. The UTMR system is able to capture practice although such possibility cannot be managed at the user level, but still at a technical one since design and implementation interventions are required. Thus, any process redefinition requires a discontinuity in the community life, since the spontaneous evolution of practice cannot be captured by its reification within technology. This circumstance makes the solution particularly rigid, and unable to easily evolve with the evolution of local practices. We believe that this gap is rooted into an assumption which characterizes KM technologies in general; according to this assumption, technology is a medium that vehicles some message and not, as proposed by many researchers, a message in itself (Orlikowski 1992). In particular, as in the case of UTMR, the workflow defines the community’s practice and, as stated by researchers on communities (Wenger 1998), practice is the major knowledge output of community’s members interactions. From this perspective, the practice (the workflow), is a knowledge product, rather than a technological premise to knowledge sharing.

Related to the previous one, another goal of the system was to respect semantic diversity and autonomy, enabling processes able to bridge semantic boundaries. But this means not just the discovery of semantic relationships among existing structures, but also the discovery of new structures that could favour the emergence of new semantic relationships. Said differently, meaning coordination is approached, but not meaning negotiation. In fact, negotiation among peers that own different conceptual schemas in order to converge to some shared ontology has not been implemented. Our impression is that this gap is rooted in a core assumption which is implicit in semantic matching technologies; that is, they assume that semantic relationships should be found among existing conceptualizations; rather, we believe that a semantic matching technology should also be able to suggest how conceptualizations should change in order to find some matching.

3. Identified Research Challenges

3.1 Ontology heterogeneity

3.1.1 Problem Statement

3.1.2 Knowledge Processing Task and Component

3.1.3 Requirements Analysis

3.2 Search for very relevant resources

3.2.1 Problem Statement

3.2.2 Knowledge Processing Task and Component

3.2.3 Requirements Analysis

3.3 Coordinated heterogeneous processes

3.3.1 Problem Statement

3.3.2 Knowledge Processing Task and Component

3.3.3 Requirements Analysis