DEPARTEMENT OF ELECTRONIC AND TELECOMMUNICATIONS

Ph.D. program in
"Telematica e Società dell'Informazione"
Cycle: XIV

Instituted by the Italian University Consortium among
University of Florence
University of Siena

MODELS AND MECHANISMS FOR SOCIAL AND BEST PRACTICE NETWORKS

A thesis submitted for the degree of
Doctorate of Philosophy

Candidate:
Michela Paolucci

Coordinator:
Prof. Dino Giuli

Supervisors:
Prof. Paolo Nesi
Prof. Alessandro Fantechi

ING-INF/05

ACADEMIC YEAR 2011/12

Greetings

I thank Prof. Giuli who offered me the opportunity to further my studies.
I thank Prof. Nesi who guided and supported me during these years.
I thank all the professors, researchers, colleagues and friends of the DISIT group with whom I got to work.
I sincerely thank my family, without which I would not even have thought of starting this path.
I would like to thank all friends and people I met and knew, people who have listened to me and who have been worried about me and welcomed me with a hug, with a word or even with a simple glance. I would like to thank those that have stimulated my curiosity, those who have amazed me, who have talked with me with sincerity and with their open-heart. I thank those that made me smile and had fun with me but also those who have challenged me, who criticized me who encouraged me to discuss and change if necessary, people who showed me new ways to test themselves.

Mchela Paolucci

Life can only be understood backwards,
but it must be lived forward.

Søren Kierkegaard

Index

List of Figures

List of Tables

Preface

Since antiquity the theme of knowledge has played an important role in the life of man. The concept of knowledge inevitably involves a multitude of disciplines. It is a very general concept that can be applied to a wide range of contexts. The concept of knowledge starts from the individual, from his critical analysis and on how one interacts with people and objects around. The goal of this thesis is to investigate the treatment of knowledge, understand which are the best technologies available on the market for its management and finally to design and implement new technologies and a framework capable of managing the knowledge flow in all its complexity and entirety.

In order to cope with the identified objectives, understand how the knowledge can be treated, investigated and then formalized and shared is fundamental. Moreover the study of the methods and models of interaction and communication among people is relevant as well. In the course of history, a wide set of models for knowledge treatment have been developed and in many of them the knowledge is not seen as a static concept but as a continuous becoming: a set of ideas that born are spread and grow up among through people interactions. One of the most relevant knowledge treatment models is the SECI, studied by Nonaka and Takeuchi, sees the knowledge as a collection of facts, information, skills that a person has acquired in the course of his/her life. The knowledge is seen as an evolving process starting from the individual and whose flow in the society: the dissemination, exchange and the reuse of knowledge generates new knowledge. Basically, two types of knowledge are outlined: explicit and tacit. Explicit knowledge is formal, systematic language, it can be expressed and shared in the form of data, in scientific formulae, systematic language, specifications, manuals and such like, it can be processed, transmitted and stored relatively easily; while tacit knowledge is that in the human mind and it is highly personal and hard to formalise, externalize or mediate, represented by using the normal channels of communication.

From the technical point of view it is noticed that the ICTs began their journey starting from the need to manage: data, information, intellectual capital within an organization and or intra-organizations. Now, they are continuing their journey influencing and interacting in the everyday life of people even and especially thanks to the Web. The organizations start to structure their intellectual capital: making use of digital documents and databases, giving importance to the classification of the knowledge and to the collaborative and group exchange of tacit and explicit knowledge; enabling voice and video communications; adopting technologies to improve the knowledge acquisition; promoting learning processes, etc. The World Wide Web has provided methods, technologies and incentives for the knowledge diffusion. It has no sense to separate aspects related to document and content management from the web or from social aspects that are fundamental for the dissemination of information and relevant to increase skills, know-how and creativity. The most advanced organizations integrate aspects: such as social and ICTs that are giving rise to flexible management systems, adaptive, scalable, multi-channel and multi mode which allow people to capture knowledge in real time in order to distribute the knowledge and to take account of its changes in updating and making it available on the basis of the needs of the organizations. However, the first

step for managing the knowledge into an organization is to acquire it, in order to make the tacit knowledge concrete in a way in which it can be quantified and treated. After that, the information collected has to be retrieved and managed in order to allow extracting new knowledge. In this flow, it is also fundamental to remember that the explicit knowledge, in its different forms, represents an intellectual capital that has to be protected. Many technologies and tools for knowledge acquisition, retrieval, discovery, extraction and protection are identified, analysed, described in details and finally compared among them. This analysis allows the identification of how has to be the best infrastructure for the knowledge treatment: which are the main technical features to be studied and implemented to take into account both the content management and social interactions.

This thesis work is organised in the following chapters:

Chapter 1 is oriented to clarify the knowledge concepts in the history, its features, the knowledge treatment models studied in course of the time, especially taking into account the SECI model, studied by Nonaka and Takeuchi, that sees the knowledge as a collection of facts, information, skills that a person has acquired in during the life. The knowledge is not seen as a static concept but as a continuous becoming: a set of ideas that born, are spread and grow up among through people interactions.

In **Chapter 2**, the technologies that can be applied to the management of knowledge flow are analysed. The most diffused ICTs on how the information can be acquired, retrieved, shared, managed, discovered, used to extract new knowledge, protected, etc. are identified, analysed, described in details and finally compared among them. As a conclusion of the comparison, a Best Practice Network is identified as the best infrastructure for the knowledge treatment taking into account both the content management and the social interactions.

In **Chapter 3**, the Architecture of the Best Practice Network developed is described. This architecture has been designed in a way in which it can be easily customised and consequently applied in many different contexts. In the subsequent chapters, three different scenarios in which the defined Best Practice Network model is applied are described.

Chapter 4 describes the Mobile Medicine health care content distribution service. It has been realised to be used for any sharing and multi-channel distribution of multimedia and digital content mainly for educational purposes and to offer support during hospital emergency occurrences and first aid treatments. In the portal, groups for discussion on procedures and content are created, thus some suggestions can be provided on both PC and Mobiles.

In **Chapter 5,** the European Collected Library of Artistic Performance Project is described. ECLAP is an ICT PsP project of the European Commission, it develops a Best Practice Network acting as a Content Aggregator, making use of advanced database and delivery tools for the production and dissemination of the rich multilingual European heritage.

In **Chapter 6,** the Sportello APRE Toscana is described. It promotes Italian and Tuscan participation in the European Union's Programmes of research,

development a Regional APRE Committee, supports, cooperates and provides the strategic policy for APRE's development activities.

1. Knowledge Management

Since antiquity the first philosophers of Western civilization, speak of man and his rational capacities. Aristotle in Metaphysics says: "all men by nature tend to know"; sentence that does understand us the theoretical character of Greek science. The Egyptian and Mesopotamian cultures developed science for eminently practical purposes, the Greeks tended to cultivate them for desire for knowledge and understanding of the why of things. The Rhind Papyrus speaks of mathematical techniques that despite their advanced nature are limited to the precept "Does always as in this case" and does not refer to the typical Greek conquest that is the formula and the law, namely the inclusion of unlimited and possible cases in a abstract rule, [Abb]. Thales, Anaximander and Anaximenes have felt the need to search for a naturalistic explanation of the world. The Pythagoreans had an intuition that will be the basis of modern science of Galileo's type: that mathematics is the most important code that can be used to interpret the reality. In Gorgias written by Plato, Socrates says that the most beautiful search of all is precisely this: investigate on which should be the man, and on what the man should do [...]; moreover he makes his the motto of the Oracle of Delphi "know thyself" and asserts that wise is only who knows that he does not know, this because only those who knows that doesn't know, tries to know. Plato says that the science which is the true knowledge, corresponds to the be, then to ideas.

The theme of knowledge, the role of reason, have offered extraordinary opportunities for reflection: starting from Cicerone: “Domina omnium et regina ratio”, (Cicerone, Tusc. Disp., 2, 21, 47); Seneca: “Ratio perfecta proprium hominis bonum est, cetera illi cum animalibus satisque communia sunt” (Seneca, Epist., 76,9) and “Natura semina nobis scientiae dedit, scientiam non dedit” (Seneca, Epist., 120, 4) until the modern times.

The concept of knowledge inevitably involves a multitude of disciplines. It is a very general concept that can be applied to a wide range of contexts. Our interest is to understand how knowledge can be treated, identified, investigated and then formalized and shared through the use of technology. The aim is to identify the best technologies that can implement the methodologies and models allowing the diffusion of knowledge.

Many different definitions can be used in the actual society to describe the **Knowledge**, it can be useful to list some of them:

- Reading some dictionary definitions:

- "A set of facts, information, and skills acquired by a person through experience or education; the theoretical or practical understanding of a subject.", [OxfDic].
- "Awareness or familiarity gained by experience of a fact or situation: the program had been developed without his knowledge", [OxfDic].
- "Human faculty resulting from interpreted information; understanding that germinates from combination of data, information, experience, and individual interpretation", [BusDic].
- "The sum of what is known and resides in the intelligence and the competence of people", [BusDic] (in an organizational context).

- Resorting to some more structured and complex definitions coming from scientific articles:
  - "Is a meaningful set of information that constitutes a justified true belief and/or an embodied technical skill", [NonKatDaS].
  - "Knowledge is a concept: you cannot see it, but can only observe its effects. Because knowledge is an invisible, intangible asset and cannot be directly observed", [Hunt].
  - "Knowledge is defined as a set of structural connectivity patterns. Its contents have proven to be viable for the achievement of goals." [MeySug].

The same process can be done for the **Knowledge Management** discipline:

- "Knowledge Management is the discipline of enabling individuals, teams and entire organizations to collectively and systematically capture, store, create, share and apply knowledge, to better achieve their objectives." [RonYou]
- "When talking about knowledge management, it is clear that we are dealing with a set of complex issues that are interrelated and cannot be segmented. This is because much of the knowledge creation activities are products of people interacting with people, people interacting with data and information, people interacting with systems and people interacting with the environment in which they operate", [SulHaw].
- "Is the process of continually managing knowledge of all kinds to meet existing and emerging needs, to identify and exploit existing knowledge assests and to develop new opportunities", [QuiLeJon].
- "the process by which organizations seek to deal with knowledge in a systematic way for creating competitive advantage", [MBra].

By analysing these definitions it is clear that the concept of knowledge starts from the individual, from his critical analysis and on how he/she interacts with people and objects around him/her. For this reason, the study of the methods and models of interaction and communication among people become fundamental. The Knowledge is not seen as a static concept but as a continuous becoming: a set of ideas that born, are spread and grow up. Knowledge management has been analysed from both theoretical and technical points of view. Many models have been

developed that are useful for understanding the dynamics and the knowledge development and at the same time a set of tools and technologies have been developed in order to manage the knowledge. In fact, as said in [SulHaw], the IT track focuses on the management of information; the knowledge management activities comprise the construction of information management systems, artificial intelligence, data mining and other enabling technologies.

In this chapter, the most popular models of Knowledge management are analysed, while the next chapter is detailed to the main diffused technologies to support them.

1.1. Knowledge modelling

The activities of creation, acquisition, management, dissemination, use and application of knowledge involved in business processes, but also in other working flows, involving individuals are heavily influenced by the dynamics of social and organizational type, [Ruff]. The Knowledge management, therefore, concerns not only economic and technological aspects but also the psychological and organizational aspects. This implies that the initiatives and activities of intellectual capital management, even if managed through adequate technologies, can not be effective and efficient when an underestimation of even one of the other aspects linked has to be verified.
In recent years, organisational aspects related to knowledge management and to intellectual capital management have been largely studied, many authors have analyzed the organizational implications related to the creation, storage, movement and use of knowledge [NonKatDaS], [NonToyKon], [MBra], [RicRic], [BenFensPer].
The knowledge management has attracted the interest of researchers and practitioners in the past decades: a large set of models for studying knowledge management have been developed. These models have been fundamentally applied on business contexts. However, as appears clear reading the next paragraphs, the methodologies adopted are so general that can easily be extended and applied in different contexts.

The model that has undoubtedly strongly influenced research in this area is the so called ‘SECI Model’ developed by Nonaka and Takeuchi starting from 1991. This model has implications both for managerial style and on organisational structure. Moreover, it has emphasised the whole human process of communication, viewing it as an essential component of organisational knowledge management and learning. The model was developed without ever forgetting that the use of technological tools can play a key role in knowledge management.
In order to understand which are the best methodologies and technical tools that can now be exploited in this context, it is essential starting from the description of the key features of the SECI model.

1.2. Seci Model (Nonaka and Takeuchi Model)

The Knowledge can be viewed as a collection of facts, information, skills that a person has acquired in the course of the life. It is an evolving process that starts

from the individual and which flows then in the society, as group of people: the dissemination, exchange and the reuse of knowledge generates new knowledge.
According to Nonaka [NonTak], there are two types of knowledge: **explicit** and **tacit**. Although highly different in nature, they are inseparable in new knowledge creation and form the epistemological dimension of knowledge, [MBra].

Explicit knowledge is formal, systematic language, it can be expressed and shared in the form of data, in scientific formulae, systematic language, specifications, manuals and such like. It can be processed, transmitted and stored relatively easily using for example [NonToyKon], [TimKuz].
The Technologies for the treatment of explicit knowledge depend on the level of structuring of the available knowledge in the organization, [Ruff]. It can occur in:

- a structured format, when stored in the database, in business process management systems, in systems for the knowledge representation that make use of ontologies;
- a semi-structured format, when stored in the web pages of corporate intranets and the internet (i.e. based on HTML and XML);
- a non-structured, when accumulated in textual documents of various kinds in the Organization

Tacit knowledge is knowledge that is in the human mind and it is is highly personal and hard to formalise, externalise or mediate; typically represented by using the normal channels of communication [TimKuz], [NonToyKon]. It is deeply rooted in action, procedures, routines, commitment, ideals, values and emotions [NonKatDaS]. It strongly depends on context.
This kind of knowledge, mainly practice, is typical of being experts in some field and can be transferred, for example, through mechanisms of imitation. It requires, therefore, specific technologies for its treatment coming mainly from communications that integrate audio, video, graphics and text [Ruff].

Although highly different in nature, Explicit and Tacit knowledge are inseparable and complementary and both essential for a new knowledge creation [MBra], [NonToyKon]. The Explicit knowledge without the Tacit insight quickly loses its meaning. Knowledge is created through interactions between tacit and explicit knowledge, rather than from tacit or explicit knowledge alone.
In the next paragraph, it is discussed the conversion of knowledge and the knowledge flow through the ontological dimensions with the SECI spiral model are be described. In the subsequent paragraphs, the other features of the SECI model are described in synthesis: definition of the concept of Ba, knowledge assets and Knowledge networking.

1.2.1. SECI: Socialization, Externalization, Combination and Internalization

As stated above, Knowledge Management is seen as an evolutionary process that starts from the individual, from tacit knowledge, and that flows in the society and in the groups of people, passing through its explicit forms. They interact each other

with and interchange in human creative activities by individuals or groups, which is the key assumption of our dynamic theory of organizational knowledge creation, [NonKatDaS].
The SECI model talks about how the tacit knowledge is converted in explicit and viceversa. There are four modes of knowledge conversion, Figure 1.1. They are: (1) socialisation (from tacit to tacit knowledge); (2) externalisation (from tacit to explicit knowledge); (3) combination (from explicit to explicit knowledge); and (4) internalisation (from explicit to tacit knowledge), [NonToyKon].
The model also embodied an interaction dynamic by which knowledge is transferred in a spiral process, allowing the knowledge value to be enhanced through exchanges between individuals and groups within the organisation [RicRic]. Through the conversion process, tacit and explicit knowledge expand in both quality and quantity.

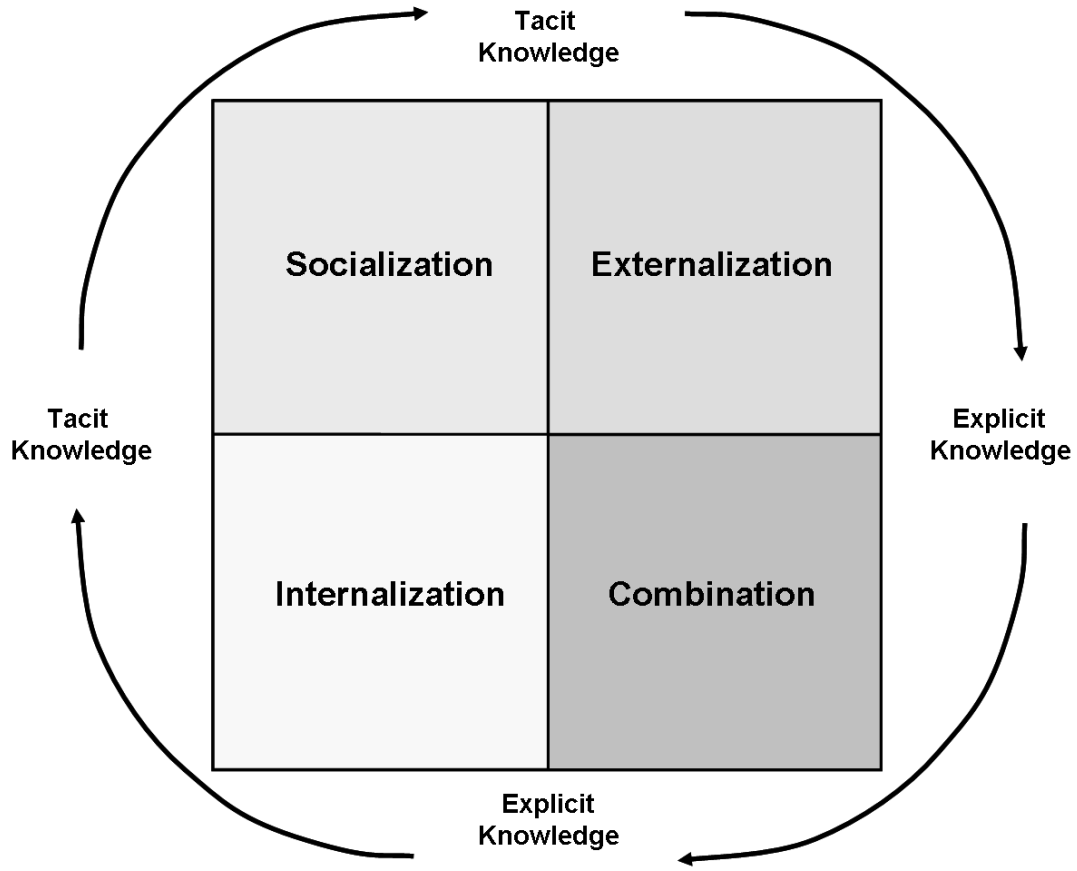

Figure 1.1: The SECI spiral model, inspired by [NonKatDaS].

Internalization: from explicit knowledge to tacit knowledge.
Internalization is a process of changing the knowledge starting from the explicit form up to the tacit form, or operational knowledge such as know-how. This transformation occurs through the internalisation: the explicit knowledge created is shared throughout an organisation and converted into tacit knowledge by individuals. This modality is triggered by the "*learning by doing or using*" concept: the individuals use explicit knowledge documented into text, manuals, sound or video formats to facilitate the internalization process. When knowledge is internalised to become part of individuals' tacit knowledge bases in the form of shared mental models or technical know-how, it becomes a valuable asset. This

tacit knowledge accumulated at the individual level can then set off a new spiral of knowledge creation when it is shared with others through socialisation [NonKatDaS], [NonToyKon].
At this mode of knowledge conversion, are particularly relevant the technology tools of: interactive e-learning, workgroup applications, technologies, synchronous and asynchronous communication (instant messaging, e-mail, chat, forums, blogs, etc.). These tools allow a two-way communication: both vertically (from teacher to learner and learner to instructor) and horizontally (from learner to learner).
E-learning and groupware are used to share with others their knowledge and educate to a continuous learning, thus developing the ability to search, select and acquire knowledge at any time [Ruff].

Socialization: from individual tacit knowledge to group tacit knowledge.
Socialisation is the process of converting individual tacit knowledge through shared experiences. Since tacit knowledge is difficult to formalise, it is necessary to pass through the shared experience, such as spending time together or living in the same environment, thereby creating common unarticulated beliefs or embodied skills. The experience can be realised in an organization in two different modalities: technical and cognitive. The first one typically occurs in traditional apprenticeships, where apprentices learn the tacit knowledge needed in their craft through hands-on experience, rather than from written manuals or textbooks. While the cognitive modality occurs beyond the boundaries of the organization: often setting up informal meetings outside the workplace, where participants have chats over sake and meals, thereby creating common tacit knowledge, [NonKatDaS], [NonToyKon]. Socialization remains always in a tacit form and therefore turns out to be very difficult to be disseminated at the organisational level. It, however, assumes fundamental importance, because it contributes to the enrichment of human capital in an organization. Socialization is supported mainly by technological tools used for cooperative and collaborative work (groupware) and for synchronous communication such as video conferencing, instant messaging, chat, e-learning. Although, each of these tools has limitations compared to the direct human relation: it is possible to say that they can increase the human socialization without to substitute the real and direct contact among people [Ruff].

Externalization: from tacit knowledge to explicit knowledge.
Externalization is the process of articulating tacit knowledge into explicit knowledge. When tacit knowledge is made explicit, knowledge is crystallised, thus allowing it to be shared with others, and it becomes the basis of new knowledge [NonKatDaS], [NonToyKon]. In order to obtain explicit knowledge, both the natural language, such as metaphors or analogies, and sketches and diagrams, images, notes, may be used. These languages allow different degrees of knowledge structuring. This facilitates the process of knowledge conversion that tends to develop around to specific patterns of communication and interaction between/among members of the organization. Externalization takes deep advantage of information technologies that enable the capture and the acquisition of explicit knowledge, its storage, according to appropriate formats characterized by different degrees of structuration and formalization. The technologies used are mainly those of knowledge acquisition, databases, document/content management [Ruff].

Combination: from separate explicit knowledge to systemic explicit knowledge.

Combination is a process of assembling, converting existing and new explicit knowledge into a more complex and systematic sets of explicit knowledge. This change is realized through social interactions among individuals and typically takes place through dialogue and communication, enabled directly in synchronous and/or asynchronous form from the information technologies. Explicit knowledge is collected from inside or outside the organisation and then combined, edited or processed to create new knowledge. The new explicit knowledge is then disseminated among the organisation's members [NonKatDaS], [NonToyKon]. It is possible to talk about the combination model for example when information coming from different groups of people or organizations, from various sources are summarized in a report, in a graph, etc. This is the area where information technologies play a fundamental role. In recent years, new technologies based on formal taxonomy and ontology representations and data management and unstructured seeds are making their appearance: these technologies are making the explicit knowledge manageable for an automatic structure. In fact, an information can be easily coded, classified, stored and retrievable. Also the technologies for the representation of workflow are making increasingly agile the formal representation of the working processes.
The most obvious consequence is that information technologies are transforming and expanding the possibilities of dealing with the structural component of the intellectual capital, that is largely composed of knowledge unsettled in information systems [Ruff].

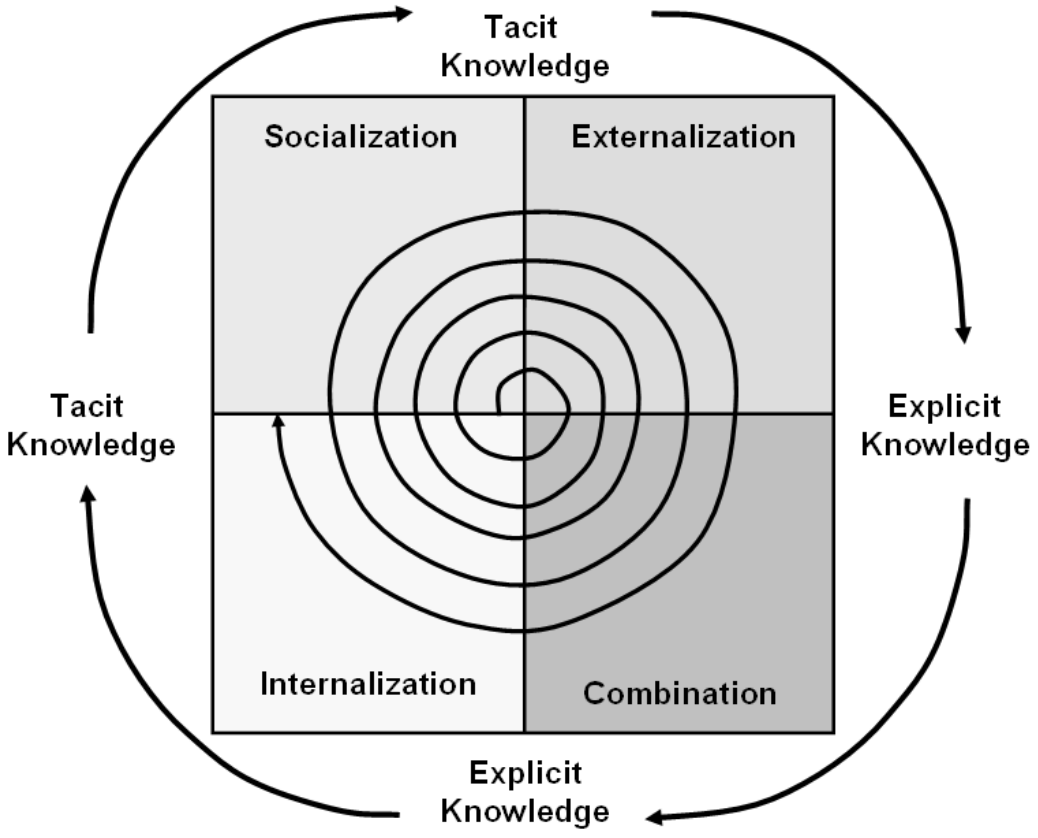

Figure 1.2: Organizational knowledge created through a Spiral.

In Figure 1.2 it is possible to see how the organizational knowledge is created through the Nonaka's spiral-like process: the knowledge is transformed through exchanges between individuals and groups within the organisation, across the four modes of knowledge conversion described above. A knowledge spiral may start from any mode, but usually begins from Socialization [NonKatDaS]. This process does not stop once it has closed the circle, but continues into a new knowledge-creating spiral, [MBra], [RicRic]. This means that the conclusion of the SECI process, which can culminate in realization of a new project or obtaining a new product, you have a new beginning of the same process but at a higher overall level of knowledge both as regards individuals that organization. This result increase the intellectual capital in its various components: organizational and relational. An appropriate organizational structure and the correct technology solutions adopted in support of such a cycle may enhance the effects of knowledge management in the growth of intellectual capital, [Ruff].

ONTOLOGICAL DIMENSIONS

Starting from the above diagrams, in which only the Epistemological dimension, relationships between the explicit and tacit knowledge, are represented, , it is possible to view a graph, Figure 1.3, in which the Nonaka's spiral is represented taking into account both the epistemological dimension and the ontological one. The figure shows that the knowledge management into an organization is a continuous stream that includes an exchange of information among individuals (groups, organizations, inter-organization) and an internal reworking of the intellectual capital (personal). According to this model, an organization is capable to create knowledge only through individuals operating within it. Therefore, the most creative people have to be valorised, stimulated and supported by offering them a creative context in which to create knowledge. The knowledge management system must, then, take care of his deposition, its circulation and its reuse so that knowledge can be stored in different repositories and treated with typical tools of information and communication technology [Ruff].

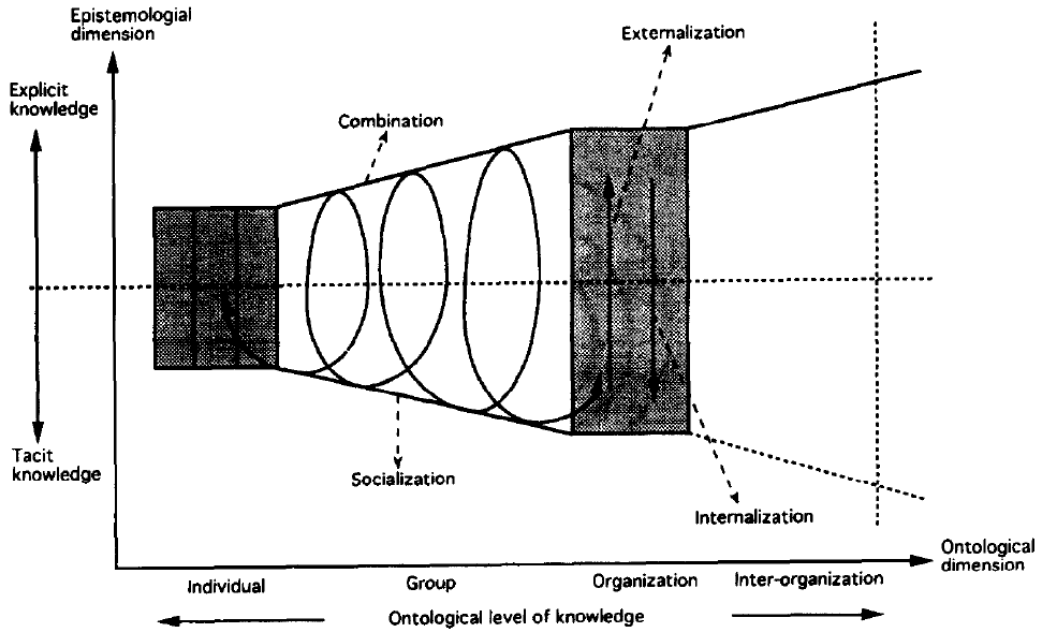

Figure 1.3: Nonaka's spiral, epistemological and ontological dimensions, source: [NonKatDaS].

1.2.2. Ba concept

Knowledge needs a context to be created as it is context specific. The context defines the participants and the nature of the participation [MBra]. The spiral of the SECI model emphasizes the dynamics of knowledge creation based on tacit/explicit conversions through the processes of socialization, externalization, combination, and internalization and happens in the so-called “Ba”: the physical or virtual place for the knowledge creation.
“Ba” is a Japanese term meaning "*place, context for exchanging creative*" proposed by Nonaka and explicit modes through which people communicate each other in order to achieve the process of conversion of tacit knowledge to explicit: it is social, cultural, and even historical, providing a basis for one to interpret information, thus creating meaning, thus becoming knowledge. The Ba can be a physical space, such as an Office, or even a shared mental model, such as that inside of a network of relationships between people who share the same interests or goals, it involves languages and communications. The Ba can also be a virtual space (Cyber Ba) created by appropriate information and communication technologies that facilitate and regulate the conversion of knowledge [Ruff]. Different types of Ba can be outlined: Originating ba, Dialoguing ba, Systemising ba, Exercising ba [RicRic].

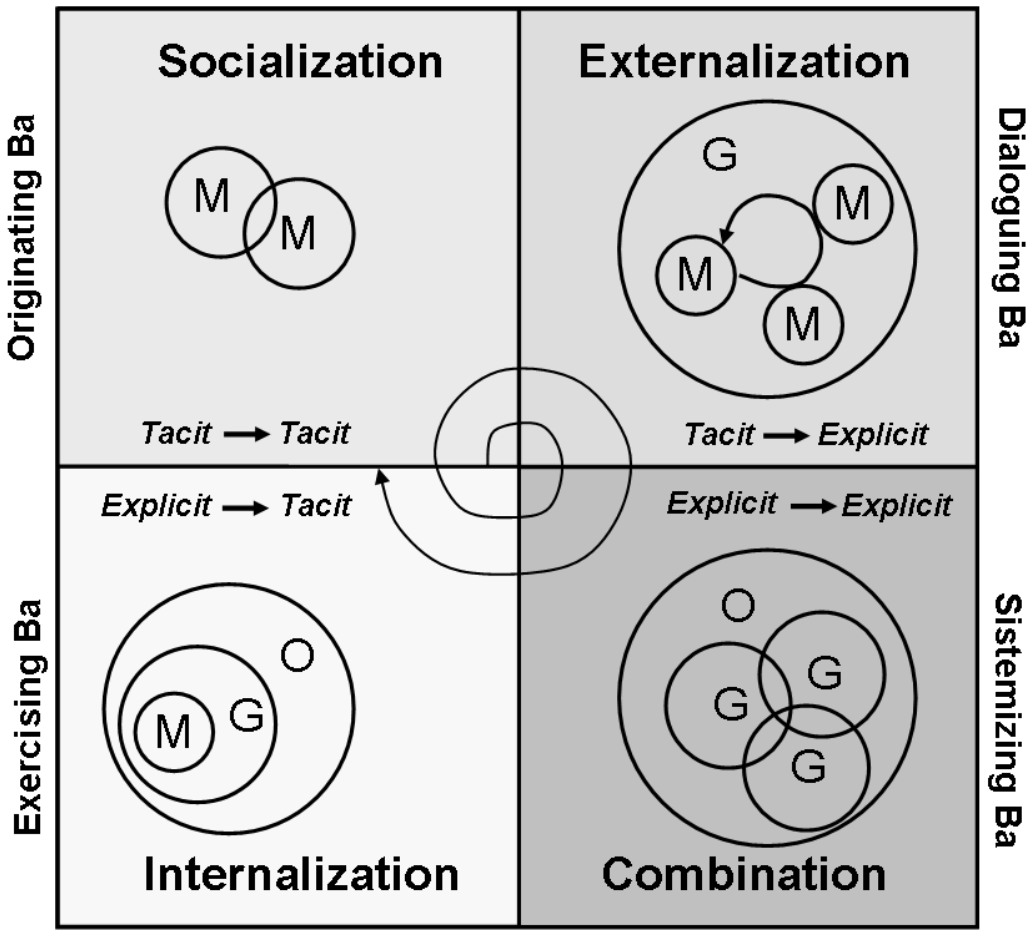

Figure 1.4: The four modes of knowledge conversion and the different types of Ba, insiperd by [RicRic].

Originating Ba

The "Originating Ba" is where the conversion from tacit to tacit knowledge (socialization) is done or when happens an exchange through face-to-face interaction among individuals. The interaction takes place in a place where you can share experiences, emotions, feelings, mental models. In order to realize this Ba is important to remove the barriers that prevent the movement of knowledge. For this purpose information and communication technologies can be helpful: multimedia communications systems, video conferencing, chat, instant messaging, etc., can simulate a possible environment for the exchange of tacit knowledge. Originating Ba is face-to-face and individual, a context for socialization: it forms the basis for knowledge conversions among individuals.

Dialoguing Ba

The "Dialoguing Ba" is the place in which the conversion of tacit knowledge into explicit (outsourcing) takes place. Is a place intended for interaction and creative discussions through which new concepts from groups or teams grow up. Community of individuals can share knowledge through physical meetings, personal tools for on-line collaborative work such as groupware, discussion forums, blogs, mailing lists, etc. Dialoguing Ba is collective and face-to-face, where knowledge is shared and converted into common terminology and concepts. Dialoguing Ba takes benefits from the participation of individuals with the right

mix of specific knowledge and when the capabilities are coordinated in a purposeful way.

Systemizing Ba
The "Systemising Ba" is the space in which the conversion, from explicit knowledge to explicit (combination), is created and enhanced by the availability of information technologies providing tools for the treatment of explicit knowledge encoded in different possible formats. Therefore, a large number of people may participate in the process of creating, systematizing and classifying knowledge. Systemizing Ba is collective and virtual. Also in this case, the information and communications technologies are effective in transmitting knowledge for example, through mailing lists, new groups, or net meetings.

Exercising Ba
The "Exercising Ba" is the context in which the conversion from explicit knowledge to tacit (internalization) takes place through on-the-job training. Also in this case the technologies information (internet, intranet, electronic manuals, expert systems, systems of Ontology-based document management, e-learning systems, etc.) support the conversion of knowledge by fostering the learning from individuals. Exercising Ba is individual and virtual, allowing for internalization of new knowledge through manuals, directories, or professional journals.

In Figure 1.4, the four different notions of Ba are defined in relation to each of the four quadrants of the SECI model that together make up the 'knowledge spiral': the M, G, and O symbols respectively represent: organization Member, Group and Organization aggregates, respectively [RicRic].

The Originating Ba can be quite spontaneous while the Dialoguing Ba is deliberate; both of them are tacit by nature. In the Systemizing Ba, the knowledge becomes more explicit, making use of forms of documentation and databanks. Moreover, the Systemizing Ba benefits from the use of information systems facilitating interaction where explicit knowledge can be efficiently exchanged and diffused through the different ontological dimensions. Finally, in the Exercising Ba, the explicit knowledge is embodied into each member of the organization to increase the know-how and become coordinated action [MBra].

1.2.3. Knowledge Assets

According to the Nonaka's definition, the knowledge assets are firm-specific resources that are indispensable to create values for the firm, [NonToyKon].
The Knowledge assets are the resources and the methodologies adopted during the knowledge creation process. They are the most important assets of an organization develop and sustain its competitive advantage. Moreover, they are constantly evolving: taking a snapshot of the knowledge assets in a specific point of time is not enough to properly evaluate and manage the knowledge asset.
To understand how knowledge assets are created, acquired and exploited, they can be categorised into four types: Experiential, Conceptual, Systemic and Routine, see Figure 1.5.

Experiential knowledge assets

Experiential knowledge assets consist of the shared tacit knowledge that is built through shared hands-on experience amongst the members of the organisation, and among the members of other organisations or customers with which it has relations. Skills and know-how acquired and accumulated by individuals through experiences at work, are examples of experiential knowledge asset. It depends from: skills and know-how acquired and accumulated by individuals through experiences at work; emotional knowledge, such as care, love and trust; physical knowledge such as facial expressions and gestures, energetic knowledge such as senses of existence, enthusiasm and tension, and rhythmic knowledge such as improvisation and entrainment.

Conceptual knowledge assets
Conceptual knowledge assets consist of explicit knowledge are articulated through images, symbols and language. They are the assets based on the concepts held by the members of the organisation. Brand equity, which is perceived by customers, and concepts or designs, which are perceived by the members of the organisation, are examples of conceptual knowledge assets.

Systemic knowledge assets
Systemic knowledge assets consist of systematised and packaged explicit knowledge, such as: specifications, manuals, etc. It is fundamental that all these intellectual properties have to be legally protected, for example using licences and patents. A characteristic of systemic knowledge assets is that they can be transferred relatively easily, for this reason the intellectual property rights play an important role.

Routine knowledge assets
Routine knowledge assets consist of the tacit knowledge that is routinised and embedded in the actions and practices of the organisation. Know-how, organisational culture and organisational routines for carrying out the day-to-day business of the organisation are examples of routine knowledge assets.

Experiential Knowledge Assets	**Conceptual Knowledge Assets**
Tacit knowledge shared through common experiences • Skills and know-how of individuals • Care, love, trust, and security • Energy, passion, and tension	Explicit knowledge articulated through images, symbols, and language • Product concepts • Design • Brand equity
Routine Knowledge Assets	**Systemic Knowledge Assets**
Tacit knowledge routinised and embedded in actions and practices • Know-how in daily operations • Organisational routines • Organisational culture	Systemised and packaged explicit knowledge • Documents, specifications, manuals • Database • Patents and licenses

Figure 1.5: Knowledge Assets features, source: [NonToyKon].

1.3. Knowledge Networking - Community/Networks of practice

In the knowledge management perspective, the term "knowledge networking" is often used. It indicates a set of: people, resources and relationships among them, who are assembled in order to accumulate and use knowledge primarily by means of knowledge creation and transfer processes, for the purpose of creating new knowledge and value, [HogEdv]. In [BaBlHe], the terms "community of practice" and "networks of practice" are used. While in [Ruff] the term used to indicate a group of people who share a common knowledge and which may have common interests and objectives is 'community of practice'. In the communities of practice, that can be constituted either formally and informally, individuals interact with each other, exchanging knowledge on topics of mutual interest; exchanging processes result in the generation and sharing of new knowledge. A community of practice can be born and grow spontaneously in any organization, by a process of socialization aimed at sharing experience and daily work practices; once formed is held together by the strong cohesion and spirit of the group of participants. The knowledge produced and traded by these social aggregations are predominantly tacit and unwritten. This because in communities of practice group membership is not questioned. An individual may belong to different communities of practice, more or less structured, and can spontaneously pass from one to another. The establishment and maintenance of communities of practice is strongly favoured by information technologies that, thanks to the tools of communication/collaboration and groupware, allow the creation of virtual communities and relocated. Information technologies allow to capture and acquire, making it explicit, a part of the tacit knowledge that is generated and exchanged within the community of practice. In modern organisations, it is increasingly widespread establishment of these communities in order to facilitate the process of creating and sharing knowledge and experiences among workers.
In Networks, many aspects, some of them connected with the SECI, are relevant [MBra]:

Purpose: it comprehends both the cognitive orientation and mental models that are assumed to guide the individual's and ultimately an organization's perception of the purpose for their activities.

Links: means the links between the different Ba's. This type of connections cannot be underestimated either and is explicated through the different types of Ba: originating, dialoguing, systemizing, and exercising.

Interactive levels: In networks, the interactions take place among individuals, among individuals and groups, among groups, and among groups and organizations, that is, in every possible way of combination across different ontological dimensions, where different types of Ba would occur.

Multiple leaders and independence of members: Since a network may consist of members in different ontological dimensions, there would be multiple leaders. Multiple leaderships are also a cultural issue. Closely tied to multiple leaderships, it is the independence of members. In shared knowledge creation, a reasonable

degree of independence of members to act should be allowed. Purpose and links/connections among people form a simple system of inputs, processes, and outputs in a network. People represent independent members, which through shared leaderships may generate integrated levels. Purpose represents cooperative goals, which are pursued through interdependent tasks generating concrete results.

A (social) network is backed up and transformed by the information and communication technologies. This network of knowledge-resources is continuously being augmented by knowledge gained, from learning situations, a knowledge network should be regarded as a dynamic structure rather than as a static institution [HogEdv].

1.4. Social Networks

The Social Network include many of the features that we have described so far: this is the reason why it has been chosen to describe them in a separate paragraph. These complex systems are web platforms often realized customizing some of the most diffused Content Management Systems (CMS) or Learning Content Management Systems (LCMS). Many widespread Social Networks (SN) are mainly focused on contents (e.g., [YouTube], [Flickr]) [KraNac], whereas others are more focused on: establishing relationships among users ([Facebook], [MySpace], [LinkedIn]), [NaRaCh]. SN analysis is typically focused on analyzing relationships among users, users and groups, in order to identify which are the most central users and groups, and, on the other hand, which are those that are frozen out, namely being those running the risk of losing interest in the network activities due to a serious lack of involvement, [BCFNPPa].

A comparison related to social networking services and the most diffused social networks available on the Web is done. The comparison includes the list of functionalities related to users and social interaction and functionalities related to content fruition. Table 1.1 and Table 1.2 summarize the results of the comparison, [EC-DE6.2.2]. These services are described and with the SECI model in Chapter 2.

Legend	
O	Object, content
U	User
Y	Yes, the service is present
N	No the service is missing
A	Audio
V	Video
D	Document
I	Image
M	Multimedia

Social Network comparison on User services					
	YouTube	**Flickr**	**Facebook**	**LinkedIn**	**MySpace**
User profile, descriptors	Y	Y	Y	Y	Y

Friends/colleagues	Y	Y	Y	Y	Y
Query on Users	N	N	Y	Y	Y
Groups and Forums	Y	Y	Y	Y	Y
Multilingual pages	Y	Y	Y	Y	Y
Invitations of users	Y	Y	Y	Y	Y
Chats, on line, messages	N	Y	Y	Y	Y
Recommendation U→U	N	N	Y	Y	Y
User Lists	N	N	Y	Y	Y
Taxonomy on Users	N	N	N	N	N
Direct call, SMS, Email	Y	Y	Y	Y	Y
Events	N	N	Y	Y	Y
Group and user statistics	N	N	N	(N)	N

Table 1.1: Social Network comparison on Users.

Social Network comparison on Content services					
	YouTube	**Flickr**	**Facebook**	**LinkedIn**	**MySpace**
Multimedia (M), crossmedia UGC	Y(M)	Y(M)	Y(M)	N	N
Audio, Video, Images, Doc	V	I, V	I, D, V	I, D	I, V
Moderated UGC	Y	N	N	N	N
Query on content	Y	Y	N	N	Y
Comments on Content	Y	Y	--	--	Y
Ranking and voting	Y	N	--	--	Y

General Recommendation O	Y	Y	Y	Y	Y
Recommendation O→U	Y	Y	--	--	Y
Recommendation O→O	Y	N	--	--	N
Taxonomy for content/profile	N	N	N	N	N
Play Lists of content	Y	N	N	N	N
RSS Feeds for content	Y	Y	N	--	Y
Links with other SN	Y	Y	Y	Y	Y
Mobile Support	N	N	Y	N	Y
DRM/CAS Support, IPR mng	Y(D)	N	N	N	N
GeoTagging and QR	N	Y	N	N	N
E-learning courses	N	N	N	N	N

Table 1.2: Social Network comparison on Content.

2. Technological tools

Since the twentieth century until today, the Information and Communication Technologies (ICTs) have developed and diffused. The ICTs began their journey starting from the need to manage: data, information, intellectual capital within an organization and or intra-organizations. Now they are continuing their journey influencing and interacting in the everyday life of people even and especially thanks to the Web.

On the one hand, the organizations start to structure their intellectual capital: in make use of digital documents and databases; in give importance to the classification of the knowledge; in the collaborative and group exchange of both tacit and explicit knowledge, in making explicit the implicit knowledge, in promoting learning processes, etc.

On the other side, the World Wide Web has provided new methods, technologies and incentives for the knowledge diffusion. Starting from the first version of the WWW in which predominantly static documents, blog, forum, simple database connections were present, passing from the web 2.0 that gives importance to the connection among the resources but also and especially to those among the people: the first communities, social networks, wiki was born in this phase. Finally, due to the excessive amount of resources that are located on the web, the society has felt the need to both categorize and classify such information in a more intelligent way, making use for example of taxonomies, ontologies, workflow, web services and also improving the potentialities and ability of search engines.

With this in mind, also in knowledge management context it has no sense to separate aspects related to document and content management from the web or from social aspects that are fundamental for the dissemination of information and relevant to increase skills, knowhow and creativity.

The most advanced organizations integrate many aspects such as: social aspects, ICTs that are giving rise to flexible management systems, adaptive, scalable, multi-channel and multi mode. These ICTs allow people to capture knowledge in real time in order to distribute the knowledge and to take account of its changes in the by updating it and making it available basing on the needs of the organization.

The aim of this thesis is exactly: to describe an architecture (without forgetting the cases study), that gives importance to all these aspects and that are described in the next chapter. However, to understand the potential offered by the different

technological solutions available is of fundamental importance for the purposes of their correct application.

2.1. Information and Communications Technologies

Many methodologies to describe the most diffused technologies for knowledge management, can be adopted. In this chapter, the ICTs are described and grouped into two main categories: mainly oriented to the content management or to the development of Social Interactions.
In the following paragraphs, these technologies are compared each other according to the following sett of factors:

- features from a technical point of view;
- kind of knowledge they most stimulate: tacit or explicit;
- which of the four modes of knowledge conversion provided by the SECI Model they correspond.

In
Figure **2.1**, the feathering serves to emphasize the relationship among the SECI model and the tacit and explicit knowledge. In the following figures, the relationship among the ICTs and the four modes of SECI model conversion is described. However for not to lose the connection with the ICTs the two types knowledge, the gradient (represented in
Figure **2.1**) will be used as background.

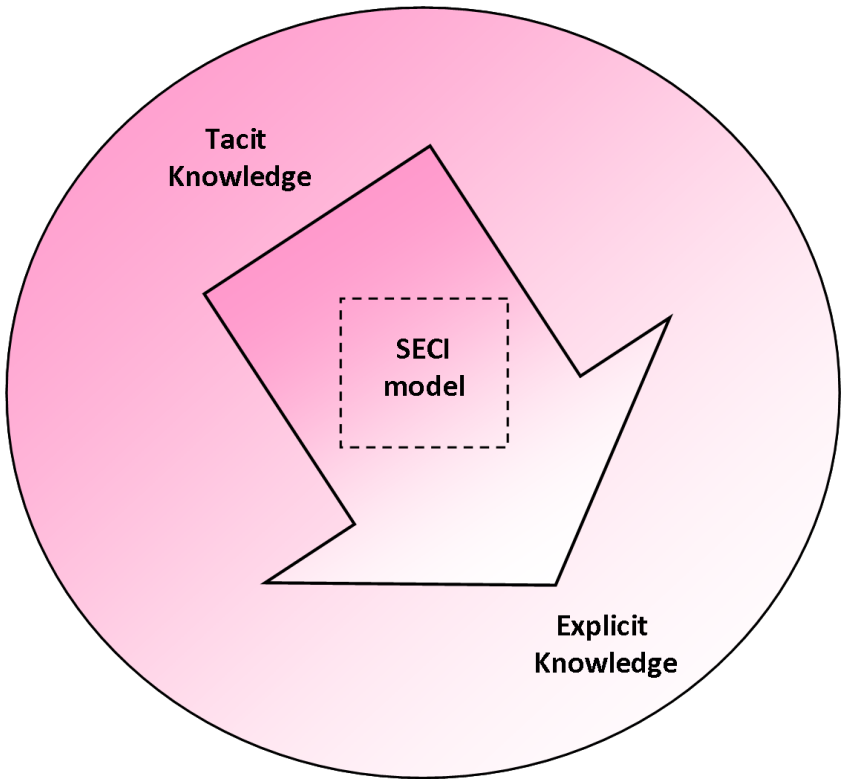

Figure 2.1: Relation among the SECI model and the kind of knowledge.

2.2. ICTs for Content Management

The first step for managing knowledge is to acquire it in order to make the tacit knowledge concrete in a way in which it can be quantified and treated. After that, the information collected has to be retrieved and managed in order to extract new knowledge. In this flow, it is also fundamental to remember that the explicit knowledge, in its different forms, represents an intellectual capital that has to be protected (e.g., through licences or patents).
To this end, it is possible to group the ICTs for content management into the following main categories:

- Knowledge Acquisition
- Knowledge Processing
- Knowledge Aggregation
- Knowledge Understanding
- Workflow

Each category is described in details in the next paragraphs.

2.2.1. Knowledge Acquisition

The knowledge acquisition can be defined as a set of methods and techniques used to concretize and extract knowledge from experts, from texts or technical manuals, from scientific works for the purpose of the transfer of the same knowledge base understandable and manageable by people from the machines. The applications or tools for the knowledge acquisition can be realized in three different formats: structured, semi-structured, non structured.

Document production
The headwaters of explicit knowledge in the form of semi-structured and non-structured format are enormously grown in number and size in recent years, this explicit knowledge is often diffused through the web, so that the internet tends to become the universal repository of human knowledge. Most of this explicit knowledge in organizations is in semi-structured or non-structured formats, in the form of: textual documents, graphics, web pages, etc. available in the intranet and on the internet.

Document production: non structured
All the formats that blend the real content of the document (logical structure) with the presentation aspects (physical structure) fall in this category. In particular, all the textual formats that store only text without giving importance to other features fit into this category. A file of this type consists of a sequence of alphanumeric characters that are displayed to the user as they appear in the file itself. The txt is

the lightest of existing formats, its key feature is the simplicity that allows it to be supported by almost all existing applications, but at the same time puts limits on it operations allowed (for example, you cannot define styles or apply formatting, etc.), [Pao].

Document production: semi- structured

Semi-structured data is a form of structured data that does not conform with the formal structure of tables in databases but nonetheless contains some markers (metadata) to separate semantic elements and enforce hierarchies of records and fields within the data. Therefore, it is also known as schema less or self-describing structure.

Metadata and metadata Standards:

This type of structure is often used for the description of the data on the web. Many semi-structured languages ([HTML], [SGML], [XML], [RDF], [OWL], [UML], etc.) exist on their bases developed also a large set of standards for the content description.

Some of the most widespread metadata standards:

- Metadata and standards for the digital content description: DC, MPEG-7, MARC, MODS, RDF, etc. ([DC], [MPEG-7], [MARC], [MODS], [RDF]);
- Metadata used in the field of cultural heritage: OAI-ORE, SBN, FRBR, etc. ([OAI-ORE], [SBN], [FRBR]);
- Metadata and languages related to the development of taxonomies and Ontologies: RFD, RDFs, OWL.

The standards' diffusion allowed to catalogue data with common languages, and thus, to enabling the interoperability between the various systems: according to this principle, for example, the Web Services have had a large development.

Multimedia content production: non textual

Generally files used for the exchanging or storing information are not linked only to writing, but also to other methodologies for the knowledge treatment such as images, audio or video. An important aspect is the ability to manipulate files in a way in which one can reduce the storage space occupied by them. This is called compression and has the goal of transforming an image (or any other file) in order to reduce the number of bytes used to scramble with consequent benefits of preservation and transformability.

Formats for images:

The simplest technique to store an image is that of the bitmap type in colour or black and white, this technique stores each pixel that composes the image. An image of this type is called aster. The major flaw of this technique is that it requires a large amount of memory that is much higher as the images are high resolution and with many shades of colour. Another storage technique is related to vectorial images that describe the individual lines and geometric figures that make up the image through mathematical algorithms. In this case, the image that has to be displayed or printed is reconstructed on the basis of the description in the file. It can be concluded therefore that the images aster are heavier and harder to edit, unlike

the vectorial images that are lighter, fully editable without losing the image quality and easy handling, [Pao]. The most diffused formats are the following:

- GIF – Graphics Interchange Format (bitmap)
- JPEG – Joint Photograph export Group
- PNG – Portable Network Graphics
- EPS – Encapsulated Postscript

Formats for audio:
In order to represent the sound in digital form, the time should be divided into intervals sufficiently small (according to Shannon's theorem); at the end of each interval is measured the level of intensity transmitted by the sound and you encode in binary format.
This procedure is called *Sampling* and is measured in: 'number of samples per second' (hertz).The quality of the conversion depends on the sampling rate, i.e., from the number of the samples that are measured each second by the signal source and from the length of the sample itself. Becomes therefore vital the study of the compression techniques that enable to obtain good reproduction of the original file but also a smaller memory footprint, [Pao]. The most diffused formats are the following:

- WAVE – Waveform audio format
- MP3 - MPEG-1/2 Audio Layer 3
- MIDI – Musical Instrument Digital Interface

Formats for audio/video:
A video consists of a series of images called frames. The frames appear one behind the other in rapid sequence. In the case of audio/video formats, the codec's task is to compress a sequence of images (video), store it in a file and then unzip this sequence to redisplay it, [Pao]. The most diffused formats are the following:

- AVI – Audio Video Interleave
- MOV – QuickTime
- MPEG Family – Moving Picture Experts Group (MPEG-1/ 2/ 4/ 7/ 21)

In Table 2.1, it is possible to see the relation among the Knowledge Acquisition modalities and methods and tacit/explicit knowledge.

Knowledge Acquisition					
Modality (acquisition modality)	Input		Method (acquisition method)	Output	
	Tacit	Explicit		Tacit	Explicit
Documents production (txt,	X	_	Formalization, coding	_	document, metadata,

doc, xml, html, etc.)					etc.
Multimedia content (audio, video, images, etc.) production (in all its different formats)	X	_	Formalization, coding, editing	_	multimedia content (respectively: audio, video, images, etc.)
(textual/video/etc.) Annotation on content	X	_	Formalization, coding, writing	_	annotation (respectively in textual, video, etc. form)
Knowledge modelling	X	_	Formalization, coding	_	ontology, taxonomy, classification, metadata, etc.

Table 2.1: Knowledge Acquisition and tacit / explicit knowledge.

In Figure 2.2 it can be seen the correspondence between the instruments of Knowledge acquisition and four modes of knowledge conversion provided by the SECI Model (described in Chapter 1).

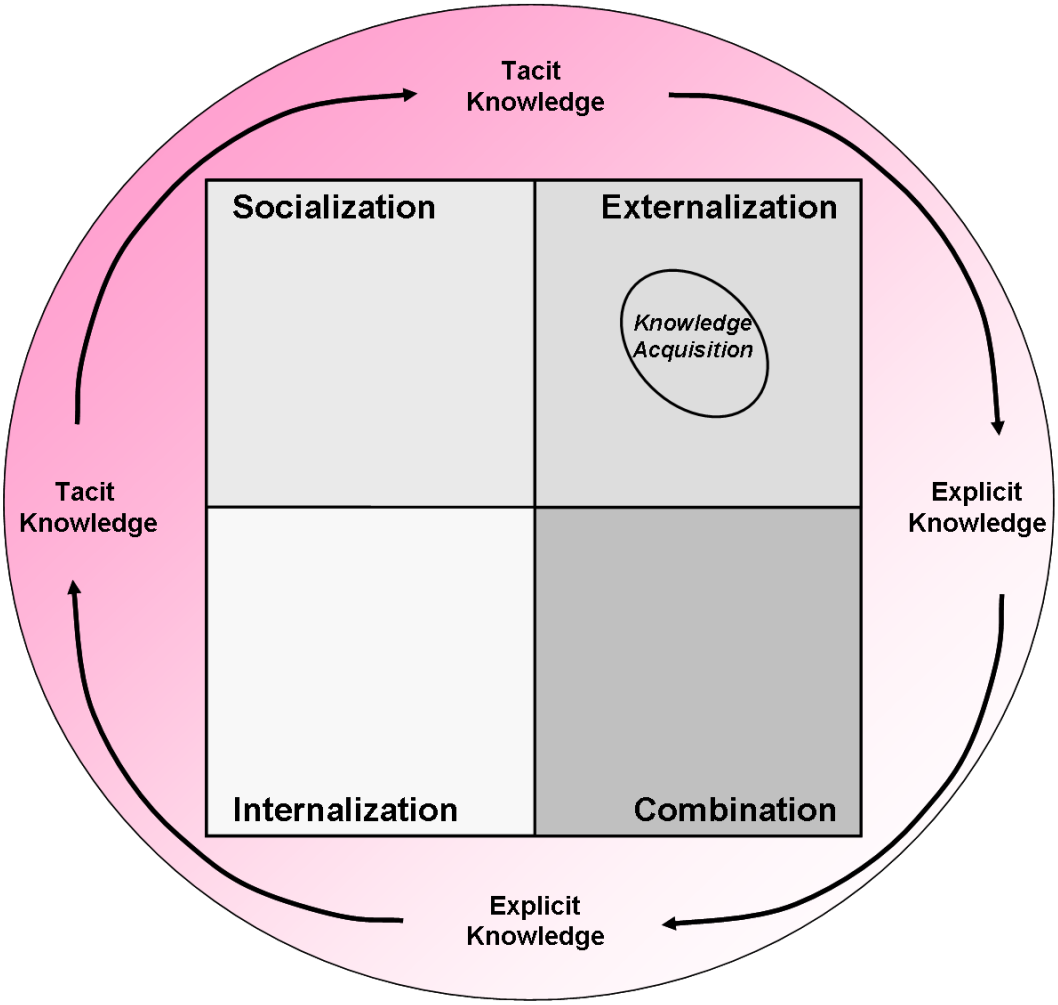

Figure 2.2: Knowledge Acquisition and SECI Model.

2.2.2. Knowledge Processing

Also the applications mainly involved for the Knowledge retrieval/ discovery / extraction can be grouped basing on the format from which they start the analysis: non structured, semi-structured, structured.

Modalities of knowledge representation

Database and data warehouse:
a database consists of a collection of structured information in which data are stored in tables, each of which describes a real world entity through a set of attributes (each row in a table represents one instance of such entities).
In relational databases, formal relations exist among tables that link together the various entities. The RDBMS (Relational Database Management System) represent one of the most common technologies used in the field of information technology.
The technology of Data Warehouse (DW) is developed next to the databases. A data warehouse, contrary to an operational database that is intended to manage many transactions (each of them is characterized by small amounts of data), is a collection of integrated data and persistent oriented to the decision support, is built

to facilitate analysis. The data in a DW come from a set of external information sources (operational databases that can be deployed over vast geographic areas and can be based on different models: relational, object oriented, lattice, etc.) and may contain useful information for decision-making processes. A data warehouse is based on the principle that the data present in an organization's operational sources must be processed and integrated in order to be effectively used within decision-making processes.

Taxonomies and Ontologies:
The Knowledge representation and the use of reasoning techniques are the basis of many applications of knowledge management. One area of science and technology that can provide an important contribution to overcoming current limitations in the formal representation of organizational knowledge is the Knowledge Representation and Reasoning (KRR).
It provides a variety of methods and techniques, mainly coming from the world of artificial intelligence, suitable to represent different forms of knowledge (uncertain, incomplete, etc.) and reasoning (deduction, abduction, planning, default reasoning, etc..) for solving complex problems. Examples of such formalisms are based on logic (datalog, disjunctive logic, logical orderly, descriptive logical, etc.) that allow a purely declarative specification of problems. In the context of knowledge representation and reasoning, the ontologies, whose purpose is to represent by means of appropriate languages and formalisms concepts resulted from an application domain, assume particular importance. The body of this formal representation is based on a conceptualization that is to specify the meaning of terms relating to an application domain, specify the conceptualization means, therefore, to attribute a meaning to unambiguous terms that define the knowledge in a specific domain.
The use of ontologies for the formal description of the conceptual structure of an application domain is an important aspect of the knowledge representation in the knowledge management. The ontologies describe classes of objects, their properties and relationships, similarly to what happens with the conceptual models for databases, by defining the map of knowledge that is provided in an application of knowledge management, [Ruff].

Methods and modalities for transforming knowledge

Natural Language Processing (NLP):
The Natural Language Processing is one of the central themes related to the management of explicit knowledge. Having tools for NLP is crucial for the treatment of the enormous amount of information available in the form of electronic text. Once that knowledge was made explicit in this form it is equally important to be able to treat it in an automatic manner, even if this task is extremely difficult.
Emerges, on the one hand, the need for tools able to interrogate the documents based on the semantics of the content, in order to capture important concepts that related to what is described within the text. On the other hand, it is fundamental take into account the language differences existing between countries that grow interest in the automatic translations.
The first step for the automatic analysis of the natural language processing is parsing that concerns the determination of the grammatical structure of a sentence. Algorithms capable of doing this exist: such as the part-of-speech tagging (POS-

tagging) through which it is possible to define the role normally played by a term (noun, adjective, verb, adverb, etc.) in a sentence. The second step is the semantic analysis which consists in identifying portions of sentences with a precise meaning mapped to the appropriate formal structures that allow the querying on the text using ad hoc languages (e.g. [Sparql], etc.).

Information Retrieval (IR):

The Information retrieval is the technology that lies at the basis of the search engines. It is grounded on indexing and consists of searching of information that have been expressed in some formal and/or natural language within text documents and web pages.

In literature, there are different approaches to information retrieval, most of them based on textual representations of documents by using templates in which the basic element for research are the words contained in the documents.

In the information retrieval, the documents are represented by appropriate models which take into account the presence of a word, and the frequency with which it manifests itself in a document, the presence of specific combinations of words representing a certain concept in a certain language. The concurrency of human language and the phenomena of polysemy, synonymy, etc., typical of the natural language, lead these tools to have limited ability to recall (the capability to found documents less than those available) and precision (the documents strongly connected to the topic sought are few with respect to the effective number into the database). The use of models and algorithmic-based techniques for natural language processing improves the performance of the engines of information retrieval.

Wrapping/Crawling:

The availability of huge amounts of information accessible through the web constitutes a substantial resource for the knowledge needs of an organization. At the same time, however, the research and exploitation of that knowledge presents several problems related to the huge size of the web and the nature of the data (structured or semi-structured) into it.

The techniques for information access on the web are generally classified into three categories: techniques based on keywords, crawling techniques and wrapping techniques.

The Crawling techniques arise from the consideration that the sets of documents available on Internet and intranets have mainly a hypertext structure that can be advantageously used for the selection of documents related to an information need. In fact, almost all documents contain, in addition to information directly represented, a set of connections (links) to other documents. And, it is quite obvious to assume that these documents contain information related in some way (from a semantic point of view) to those contained in the source document. This strategy provides, on average, a better precision than strategies based only on keywords.

The idea of wrapping techniques is to extract data and information from the documents, also using the semi-structured languages offered by web (HTML, XML), to convert the documents from the non structured or semi-structured format in a structured format. This allows you to apply to the documents, so transformed, advanced query techniques, such as those in use in the management of databases, which ensure high standards of accuracy and completeness. The wrapping creates, therefore, an image of structured documents, for example, in the form of a XML document well structured or a database table that can then be analyzed by

appropriate techniques. The wrapping techniques applied in a visual way to documents in any format, and/or to the textual output of legacy applications, are known as web scraping. These techniques are used to integrate legacy applications into new design and technology architectures to acquire data from sources of textual information.
New families of technologies for crawling and wraps are being born from the merger of these technologies with models and methods derived from information extraction and natural language processing [Ruff].

Knowledge Discovery in text:
The Knowledge Discovery in text is the result of a set of techniques that derive from those used in the knowledge discovery in databases (KDD) applied to information sources available in textual formats. The Knowledge discovery is strongly linked to the data acquisition techniques. It is typically divided into the following stages: pre-processing document acquisition, text mining, result interpretation and refinement.
The purpose of these technologies is to extract automatically the knowledge from large documentary bases through, for example, the classification of documents according to subject matter; the grouping of documents of written by the same author or related to the same topic, etc. During the document acquisition, the information is acquired through techniques and algorithms of crawling, collections of documents of potential interest in various formats, coming from different sources (web, intranet, databases, text, etc.). These documents, usually referred to a same standard format, are stored in a repository.
In the document *pre-processing* phase, each document is parsed to extract the features that characterize it on the basis of which is it mining. In this way the documents stored in the repository have a structured form dependent on the nature of the extracted features.
The type of features depends on both mining algorithms that are used to examine and also from the type and form of knowledge that you want to extract. The extractor features depend on both the basic technologies used (e.g. stemmer, regular expressions, POS-tagger-English Dictionary, etc.) and on the precision and comprehensiveness that can guarantee the extraction process.
The *text-mining* phase consists in applying a set of methods, techniques and tools for discovering regularities of information sources within semi-structured or unstructured formats. This step is closely related with the pre-processing phase that should prepare information suitable for text-mining algorithms by extracting features.
If the pre-processing phase, reaches a high level of structuring, the techniques usable for text mining are not necessarily distinguishable from those of data mining. Therefore, classification algorithms, clustering and generation of rules can be usefully applied to the representation (structured) texts produced by the pre-processing phase.
In *the results interpretation and refinement* phase, the knowledge extracted is displayed: it can occur in various forms (groups of documents with similar content, lists of concepts contained in the documents, associations between documents, temporal trend on content of documents, etc.). The display can also occur after refining processes, that have been made through special interfaces or automatic modules capable of showing the final results according to user's requirements.

Database and Data warehouse querying:

The relational databases allow not only the creation of tables, but also the manipulating of tables and of the data within them. The simplest operation that can be performed on a database is the *query*: this operation allows the user to retrieve data from the database tables, saying that the recovered data must meet certain criteria. The most diffused language to realize the queries is the Structured Query Language (SQL).

However, in large data warehouse environments, many different types of analysis can occur. In addition to SQL queries, more advanced analytical operations on data can be applied. Two major types of such analysis are OLAP (On-Line Analytic Processing) and data mining.

The Queries on the db return just a series of data that meets a set of constraints, while the *data mining* makes use of data models. These models can be seen as a synthesis of all underlying data and are in most cases more useful raw data. The mining models can take many different forms, including: decision trees, decision rules, clustering, etc.

The OLAP techniques have been studied to analyse the data contained in a data warehouse. The OLAP tools designate a set of techniques for interactive and fast analysis of a large amounts of data, which can be examinated in rather complex ways. This technique consist in starting from a set of data in a given time (such as those contained in a relational database), to transform these individual information in multidimensional data, and subsequently to run queries on the data well structured. In this way you can get answers in a reduced time.

These three tools complement each other and allow developers to pick the tool that is more right for their application

Knowledge Discovery in databases:

The Knowledge Discovery is the result of a set of techniques that derive from those used in the knowledge discovery in databases (KDD) applied to information sources available in textual formats. The knowledge discovery process in databases, fully defined by Feldman in 1996, is characterized by a structure, now finally accepted by the scientific community, composed of four phases: acquisition document, document pre-processing, text mining, result interpretation and refinement.

It should be stressed that the role of knowledge discovery in databases is not reduced only to the resolution of the important problem of selection of documents relevant to a given information need. In fact, it can potentially play an important role also to address the more general problem of the content management within organizations. An important example of such a broad potential application is the possibility to realize, through text mining techniques, tools for automatic classification of document content, useful for the implementation of various services for using the contents.

Content Based Filtering and Decision support systems:

A Decision Support System (DSS) is a decision support software that allows the users to increase the effectiveness of the analysis as it provides support to all those who must take strategic decisions in the face of problems. The primary function of a DSS is to extract in a short time and so versatile, useful information for decision making, coming from a significant amount of data. This support can be realised under many different forms: lists of suggestions and recommendations, questions on specific tasks, etc.

A DSS extracts the required data to support users directly from a database or relies on the results obtained through the Content Based Filtering techniques. These techniques often include a series of algorithms to determine the groupings of users and/or resources basing on the concept of distance between two or more items:

- Similarities and distances between content and users, branches of concepts, etc.
- Connection/Aggregation/Composition of contents,
- Statistical analysis, empirical analysis, clustering
- Automatic calculation: automatic adaptations of contents, routes, optimization, etc.
- Deductions/Inferences in the knowledge base
- etc.

Content sharing:

One of the fundamental aspects to facilitate collaborative work, in addition to the tools that facilitate the communication and the exchange of ideas among users, is the ability to share resources and develop new resources always acting in groups or working synchronously on the same resource. For the resource sharing, appropriate rights of access to the resources can be established, instead specific tools exist to regulate the collaborative editing: Concurrent Versioning System (CVS), Subversion (SVN), etc. Some mechanisms or tools that can help people to produce content in shared environments are : structured repositories, well organized file management systems, knowledge maps or semantic analysis of the content and users involved in an organization, use of white-boards, user profiling, etc.

Knowledge protection

Taking into account the emergence of distributed systems, one of the biggest problems that have to be treated for the management of digital contents and for the legal protection of the intellectual properties is that linked to problems of privacy (right of rectification of personal data, the right to cancellation of its own resources, the need to avoid identity theft, etc.) and to the copyrights on the content. This problem can be solved starting from the access control policies, until the appropriate management of access rights to the contents (DRM-Digital Rights Management) such as using licences associated to the digital contents. Knowledge protection can be regarded as the formalization of knowledge IPR (intellectually property right). This means to make explicit the information regarding the knowledge ownership. Here after some details regarding the most diffused technologies and licences:

- *Digital Rights Management (DRM):* Indicates technological systems through which holders of copyright (and so-called neighbouring rights) may exercise and administer their rights in the digital environment, thanks to the possibility of making protected, identifiable and traceable the works of which they are authors or of which they have the copyright. By using DRM, audio or video files are encoded and encrypted to ensure a more difficult spread, impediments for the users and in order to allow a use that can be: limited (e.g. only for certain periods of time or for certain uses); or

predefined in the access licence provided (separately) to the users. The files thus products carry the endorsement of copyright, and can be enriched with other information, such as pictures, biography, links, etc. The access to the content by end-users is done according to procedure profiling and authentication that allow you to distribute files required in the rules of the license signed by the user.

- *Licences:* When the copyright holder (the author but also as a person who purchased them) releases his work can bind to such a use license that describes all the actions that can be performed by third persons as regards the use of the work. A copyright holder can use A licence can be written ex-novo or the copyright holder can associate to the work one of the most diffused licences:
    - GNU licences:
        - Free Documentation License (GFDL), a license for free content;
        - GNU GPL o LGPL, licenses used for the software;
        - etc.
    - Creative Commons Licences: a set of licences that allows those who hold copyright rights to broadcast some of these rights to the public and keep the other, by means of a variety of licensing schemes and contracts that include the target of a private good to the public domain or license terms of open content (open content);
    - etc.

In **Table 2.1**Table 2.2, it is possible to see the relation among the Knowledge Processing modalities and methods and tacit/explicit knowledge.

Knowledge Processing					
Modality (procedure followed)	Input		Method (processing method)	Output	
	Tacit	Explicit		Tacit	Explicit
Cataloguing	_	Metadata	Ingestion, inputting, indexing, etc.	_	Metadata on Data Base (DB), RDF store
Knowledge Retrieval	_	DB, RDF store	Search, browse, Information Retrieval, Knowledge Discovery in DB, Database and Data warehouse querying, etc.	_	DB subset, Data subset with or without automated inference

Knowledge discovery	_	Set of DB, RDF store	Search, retrieve, Knowledge Discovery in DB/text, Database and Data warehouse querying, etc.	_	Subset of the DB set
Knowledge extraction	_	documents, DB, multimedia content (audio, video, images, etc.), etc.	Natural Language Processing, multimedia content (audio, video, images, etc.), processing, NLP Documents, etc.	_	Metadata, descriptors, Ontology, etc.
Knowledge protection	_	Metadata, licence, knowledge, content, etc.	Tools for Licence/ statement, production	_	Content accompanied by: licence, statement, protected content, etc.
Distribution	_	content	Download, streaming	_	Content
Access	_	content	Play / Read	_	View
Aggregation	_	content	Wrapping, packaging, crawling, etc.	_	Aggregated content, packaged content
content sharing	_	X	content managing / sharing tool	_	X
Decision Support System	_	Knowledge set	Inference	_	Enriched knowledge set, learned knowledge
Content Based Filtering	_	Knowledge set	Inference	_	Filtered knowledge set

Table 2.2: Knowledge Processing and tacit / explicit knowledge.

In Figure 2.3, it can be seen the correspondence between the instruments of Knowledge Processing and the four modes of knowledge conversion provided by the SECI Model (described in Chapter 1).

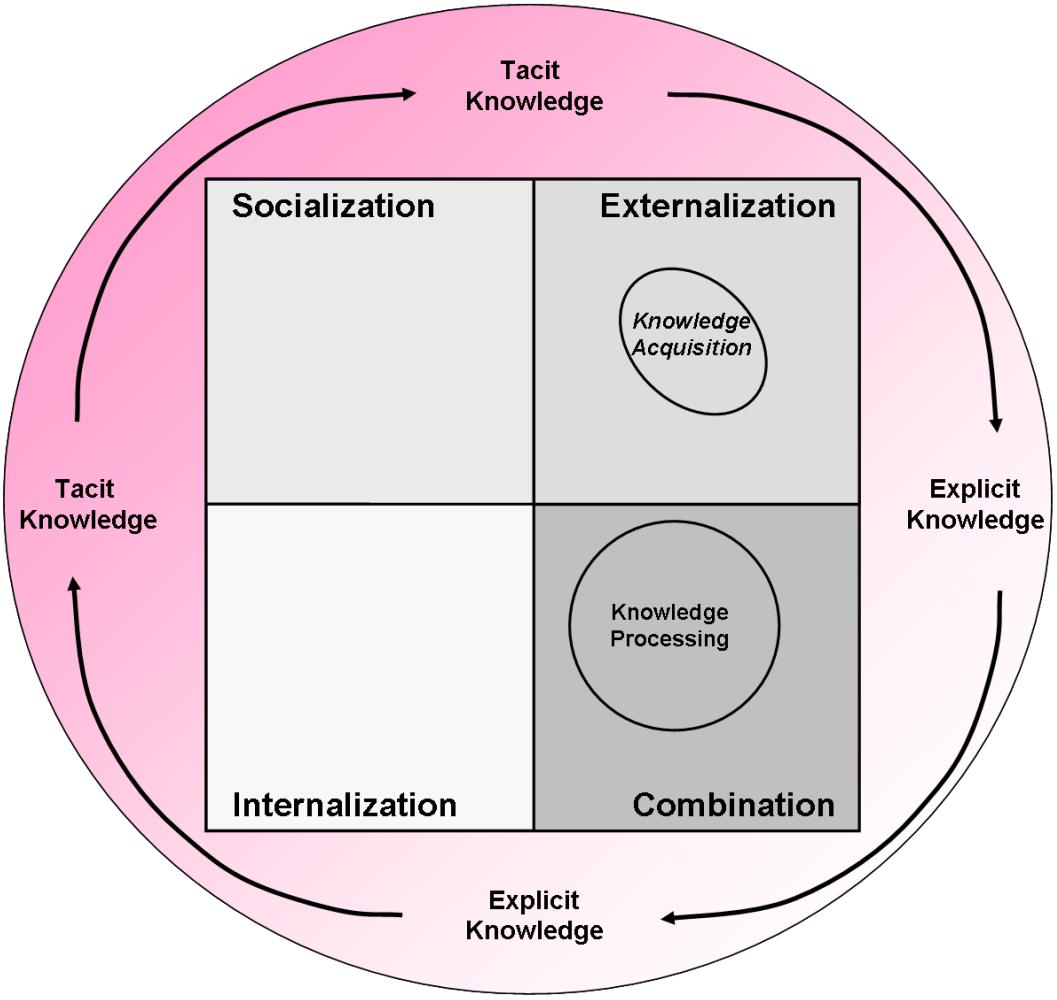

Figure 2.3: Knowledge Processing and SECI Model.

2.2.3. Knowledge Aggregation

The Information and Communication Technologies mainly oriented towards the development of Social Interactions belong to a category of software applications for knowledge management rather broad: it includes tools that help individuals to communicate and work together in groups and teams of different size and that can be geographically distributed.
This type of technology aims to encourage and facilitate the work of the Group and socialization also at distance through enabling synchronous communication (such as instant messaging, that leverages peer-to-peer paradigm, chat, video conferencing, voice over IP, etc.) and asynchronous (e-mail, forums, blogs, etc.), the use of shared documents, etc. Thanks to these technologies, knowledge is disseminated not only among groups of people who are in the same physical space but also between geographically distributed individuals who can share experiences and work together. In fact, the electronic communications can be properly captured, scanned, classified and managed for the purposes of wider knowledge management purposes.
It is possible to distinguish the following categories:

- Groupware and communication and collaboration tools
- e-learning

Groupware and communication and collaboration tools
With groupware and communication and collaboration tools can be exchanged: experiences, impressions, and opinions, to meetings; it is possible attend presentations, have discussions on many different topics, share documents, produce documents in a collaborative, etc. These tools enable people to work collaboratively through the creation of special virtual work environments and direct channels of communication among individuals or groups. Communications can be:

- Synchronous: the users involved in the communication are connected (to internet or to an intranet) at the same time range and can interact simultaneously. For example: instant messaging, video conferencing, etc.
- Asynchronous: the communication happens when the users are connected (to internet or to an intranet), not necessarily at the same time. For example: e-mail, forum, blog, etc.
- Multi-channel: i.e. able to exploit different communication channels such as networks, telephone networks, etc.
- Multimodal: i.e. usable with different devices such as computers, mobile phones, PDAs, haptic device, voice, weight, hand gesture, body posture and movements, etc.

e-learning
The e-learning technologies have a great impact on the creation, maintenance and management of intellectual capital in its human component allowing the creation of value through a process of diffusion of knowledge that derives from a combination of knowledge management technologies typical of innovative pedagogical models.

The e-learning tools allow instant access to educational materials available (anywhere / anytime) are based on the philosophy of "learning by doing" that implies: the integration of training with the daily work, the constant monitoring of the level knowledge gained on a theme, forms of cooperative learning, until the creation of communities of practice made up of individuals with common interests.
The e-learning applications are benefiting from the integration with technologies such as knowledge discovery and ontologies that allow the profiling of users based on their training needs and a more efficient organization of content and teaching activities. In this way it is possible to classify an organization's human capital in a more careful and thorough. This facilitates the orientation of the growth in the direction indicated from the strategies of knowledge management and intellectual capital.
These tools allow a two-way communication: both vertically (from teacher to learner and learner to instructor) and horizontally (from learner to learner). E-learning and groupware are used to share with others their knowledge and educate to a continuous learning, thus developing the ability to search, select and acquire knowledge at any time, [Ruff].

These technologies are capable of “channelling” the tacit knowledge, transfer conditions recreating typical direct interactions between people. Another important task performed by these technologies is to capture the implicit components of

knowledge possessed by individuals making explicitation, sedimentation and the circulation in your organization. Multimedia communications, audio, video and text can be captured, made explicit and classified in the form of electronic text or database by using the tools of acquisition described in the preceding paragraphs, [Ruff].
Below is a brief description of the main **modalities** of communication, useful to manage the Social Interactions.

e-mail:
This is the method of communication now almost old-fashioned and simple but which still remains the most popular. It is mainly used for private communications among two or more users. Therefore does not aim to encourage socialisation between individuals who do not know each other but rather of consolidating knowledge with people already known. It is a tool for sending communications, messages or private documents. In case the communicationsto be considered legally valid, the certified e-mail has to be used.

Mailing lists:
A mailing list is a collection of names and addresses used by an individual or an organization to send material to multiple recipients with a single email sending. The term is often extended to include the people subscribed to such a list, so the group of subscribers is referred to as "the mailing list". In some cases, the sending of a message on the mailing list is moderated by a coordinator, that may not approve a new posting.

Blog:
It's a website generally maintained by a person or an institution in which the authors or author publishes more or less sporadically some thoughts, news, opinions, thoughts, observations, along with possibly other types of electronic material such as images or videos. Facilitates the expressivity and the writing skills, the critical thinking and thoughtful but also and above all, dialogue and collaborative capabilities, [Cin].
Although the blog is a tool created as an expression of the individual operating modes it certainly is useful for the creation of a social network. In those cases, the blog is frequently open to all registered users, as a sort of common chat, of shared whiteboard.

Forum:
An Internet forum is an online discussion site where people can hold conversations in the form of posted messages. The messages are at least temporarily archived and accessible for other users in the future. Also, depending on the access level of a user or the forum set-up, a posted message might need to be approved by a moderator before it becomes visible.

Wiki:
A Wiki is a website which can be edited by its users and whose contents are developed collaboratively by all those who have access. The use of wiki strengthens the capability of self assessment related to the work process; the work of the learning communities enables a quick reconnaissance of each job step and the co-construction of the knowledge. The focus is entirely focused on people, then on aspects of socialization and collaboration. The wiki integrates different media:

video, audio, text, hypertext, images, etc., [Cin].There are other forms of WiKi for example the Google Docs.

Instant Messaging/chat/video chat:
Online chat may refer to any kind of communication over the Internet, that offers an instantaneous transmission of text-based messages from sender to receiver. The online chat may address as well point-to-point communications as well as multicast communications from one sender to many receivers. The video chat is established among individuals using not only text but also an audio video connection. These tools have properties intermediate between the direct conversations, e-mail and the telephone conversations because they are a less invasive way of being interrupted by a person who wants to start a conversation, but more effective and complete than using a call because it allows the simultaneous communication with multiple people to which, for example, may be submitted a question to which everyone can respond.

In Table 2.3**Errore. L'origine riferimento non è stata trovata.** it is possible to see the most knowledge Aggregation tools in relation with tacit and explicit knowledge.

Knowledge Aggregation Production (write)					
Modality	Input		Method (application/ tool)	Output	
	Tacit	Explicit		Tacit	Explicit
e-mail, mailing list	40%	60%	e-mail archive and accessing tool	_	X
Blog	50%	50%	blogs on platforms	_	X
Forum	80%	20%	forum on platforms	_	X
Wiki	80%	20%	wiki on platforms	_	X
Instant messaging / chat / video chat	80%	20%	chat tools	_	X
Lesson, questionnaire, educational game, etc.	70%	30%	e-learning activities	70%	30%

Table 2.3: Knowledge Aggregation and tacit / explicit knowledge.

In Figure 2.4 it can be seen the correspondence among the knowledge Aggregation and the four modes of knowledge conversion provided by the SECI Model (described in Chapter 1).

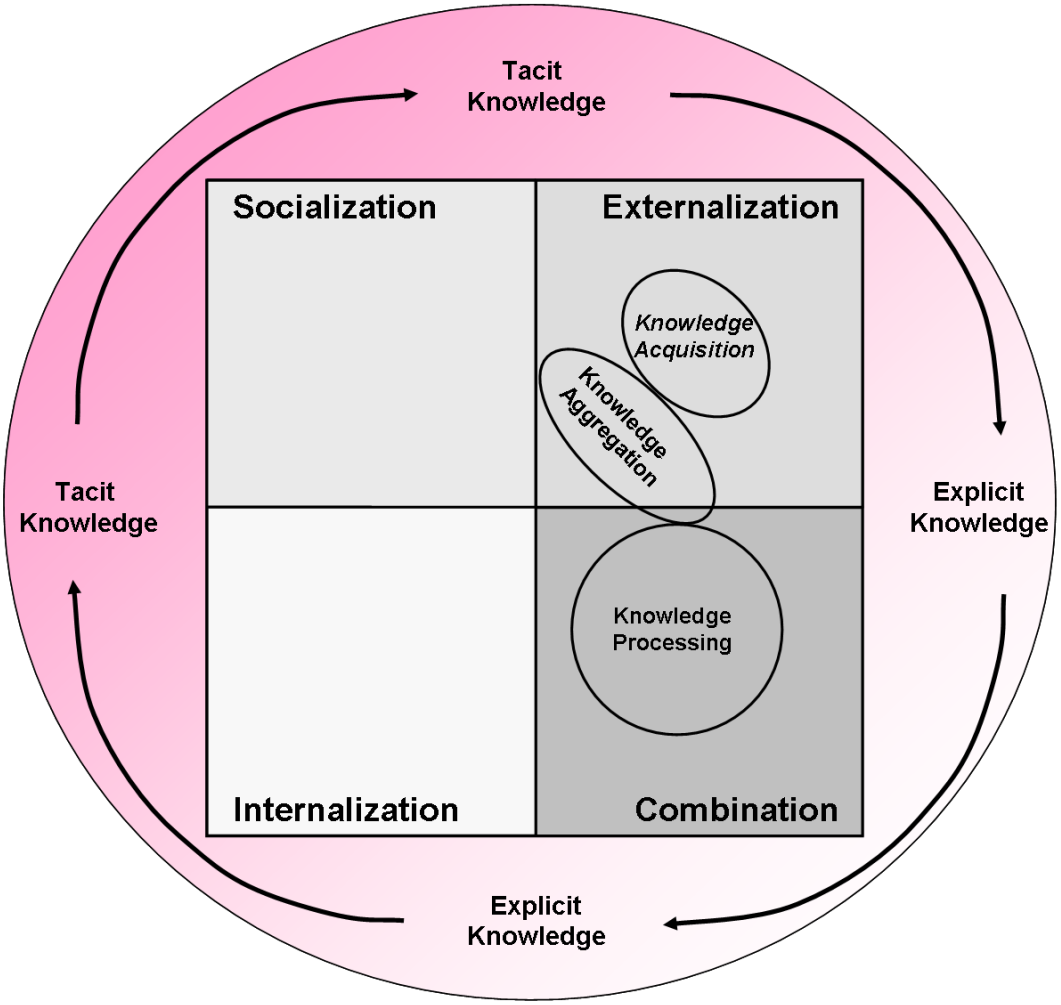

Figure 2.4: Knowledge Aggregation and SECI model.

2.2.4. Knowledge live informal interactions

In this category it is described only the VoIP, this tool allows the users to exchange knowledge without technology mediation: the knowledge flow is from tacit to tacit.

Voice over IP (VoIP):
Voice over IP (voice over Internet Protocol), or VoIP, is a technology that makes it possible to make a phone call using an Internet connection or any other network dedicated to packet switching using the connectionless IP protocol for transporting data.

In Table 2.4,**Errore. L'origine riferimento non è stata trovata.** it is possible to see the knowledge live informal interactions in relation with tacit and explicit knowledge.

Knowledge live informal interactions					
Modality	Input		Method (application/ tool)	Output	
	Tacit	Explicit		Tacit	Explicit
VoIP	X	_	VoIP tools	X	_

Table 2.4: VoIP and tacit / explicit knowledge.

In Figure 2.5, it can be seen the correspondence among the Knowledge live informal interactions and the four modes of knowledge conversion provided by the SECI Model (described in Chapter 1).

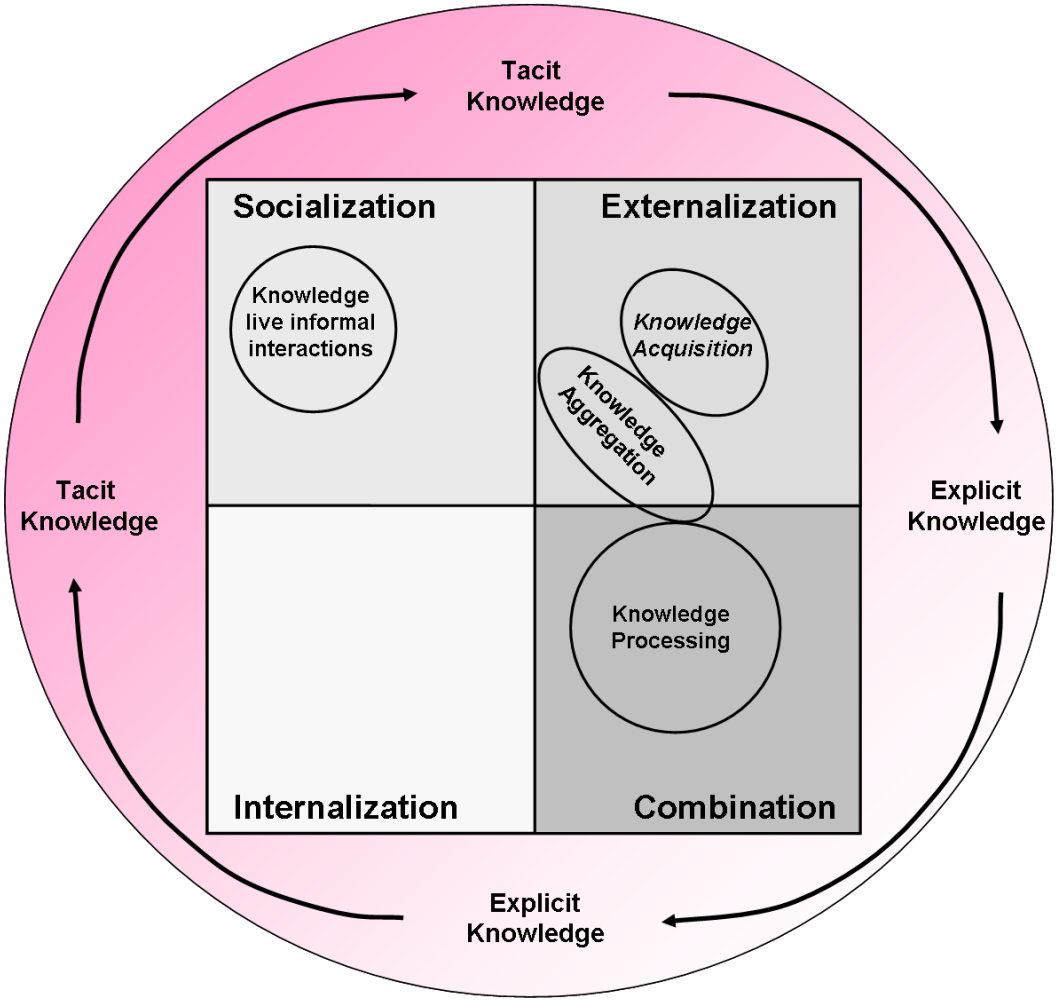

Figure 2.5: Knowledge live informal interactions.

2.2.5. Knowledge Understanding

This kind of knowledge occurs when users access to the content, display information and rework on it. In Table 2.5,**Errore. L'origine riferimento non è stata trovata.** it is possible to see the knowledge live informal interactions in relation with tacit and explicit knowledge.

Knowledge Understanding					
Modality	Input		Method	Output	
	Tacit	Explicit		Tacit	Explicit
Read	_	Document (e.g: text, mail, blogs, wiki, etc.)	Human, Natural Language Processing, etc.	X	_
Play	_	Multimedia content (image, audio, video, etc.)	Tools for playing content	X	_
Recommendations Access	_	Recommendations, suggestions	Tools for elaborate and list the recommendations	X	_

Table 2.5: Knowledge Understanding and tacit / explicit knowledge.

In Figure 2.6, it can be seen the correspondence among the Knowledge Understanding and the four modes of knowledge conversion provided by the SECI Model (described in Chapter 1).

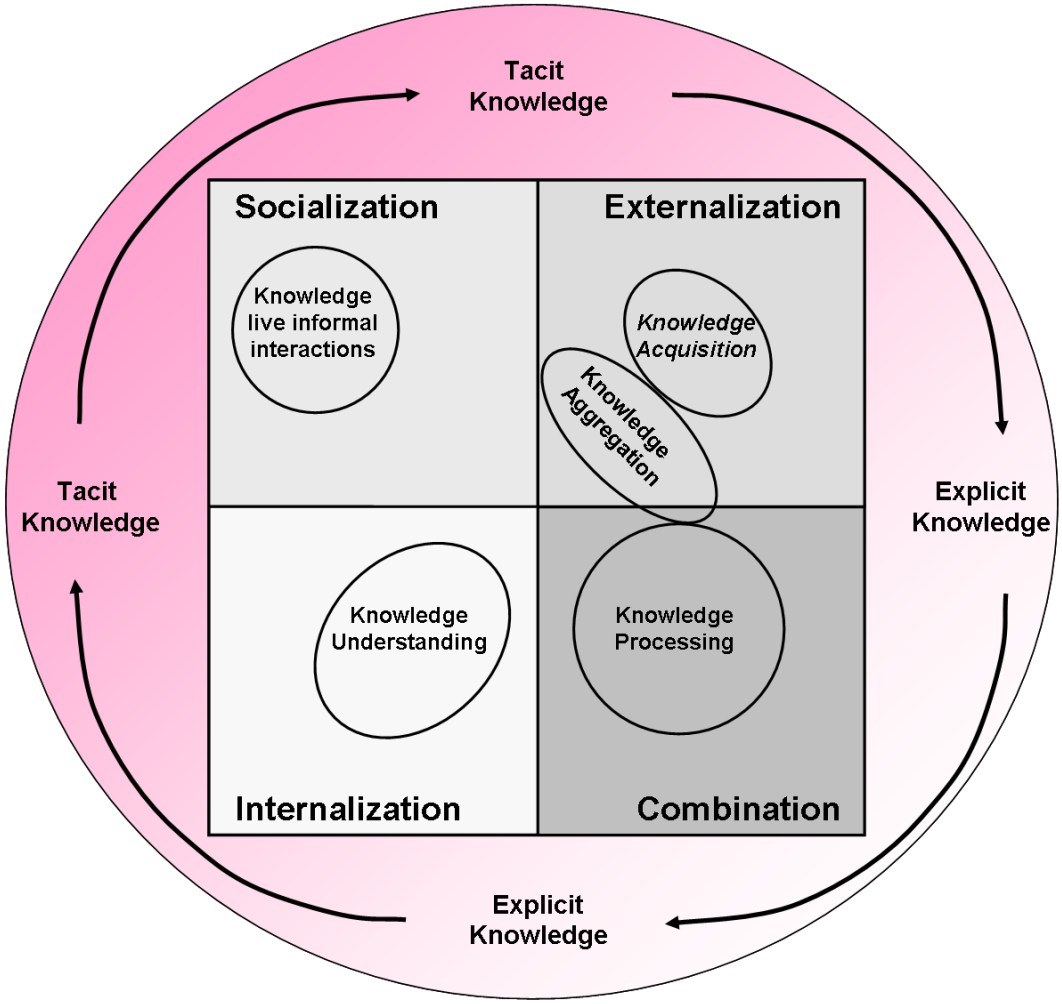

Figure 2.6: Knowledge Understanding and SECI model.

2.2.6. Workflow

Workflow includes a set of activities involving both automatic procedure (explicit knowledge treatment) without forgetting the skills and the tasks of the people involved in the processes that have to be managed. Some workflow systems, here after described as using intelligent workflow, provide to the users the possibility to choose the actions that have to be done: in this case is present a flow from tacit knowledge to explicit. It involves all the knowledge types that have been previously described (this is the reason why it has been placed in a separated section).

The workflow management system (WFMS) have become a key technology to manage business processes in many application domains such as industrial production, cash management, streamlining of information processes, monitoring of hospital activities, etc.
According to the Workflow Management Coalition (WfMC), "Workflow is concerned with the automation of procedures where documents, information or tasks are passed between participants according to a defined set of rules to achieve, or contribute to, an overall business goal. Whilst workflow may be manually organised, in practice most workflow is normally organised within the context of an

IT system to provide computerised support for the procedural automation and it is to this area that the work of the Coalition is directed.". To this end, Workflow Management Systems (WMS) have been developed. A WMS "provides procedural automation of a business process by management of the sequence of work activities and the invocation of appropriate human and/or IT resources associated with the various activity steps", [Holl]. Each workflow is composed of a number of activities and sub-processes (procedures) that have to be followed to obtain a certain result from certain initial conditions.
The intelligent workflow represents an evolution of workflow management systems that integrate these tools within typical knowledge management, it is able to implement a policy of "intelligent" management of business processes. These technologies enable the construction of communication infrastructures by which the various actors that contribute to a process can communicate with each other, for this reason the technologies used can be considered as an integral element of the architecture, [Ruff].

In Table 2.6,**Errore. L'origine riferimento non è stata trovata.** it is possible to see the knowledge live informal interactions in relation with tacit and explicit knowledge.

Workflow					
Modality	Input		Method (type of knowledge involved)	Output	
	Tacit	Explicit		Tacit	Explicit
workflow (as a set of automatic procedures)	X	_	A set of procedures in: Knowledge Acquisition	_	X
	_	X	A set of procedures in: knowledge Processing	_	X
intelligent workflow	X	_	A set of procedures in: Knowledge Acquisition	_	X
	_	X	A set of procedures in: Knowledge Processing	_	X
	_	X	A set of procedures in: Knowledge Understanding	X	_
	X(%)	X (%)	A set of		X

			procedures (also providing questions to the users, etc.) in: knowledge aggregation		
	X		Knowledge live informal interactions	X	

Table 2.6: Workflow and tacit / explicit knowledge.

In Figure 2.7, it can be seen the correspondence among the workflow and the four modes of knowledge conversion provided by the SECI Model (described in Chapter 1).

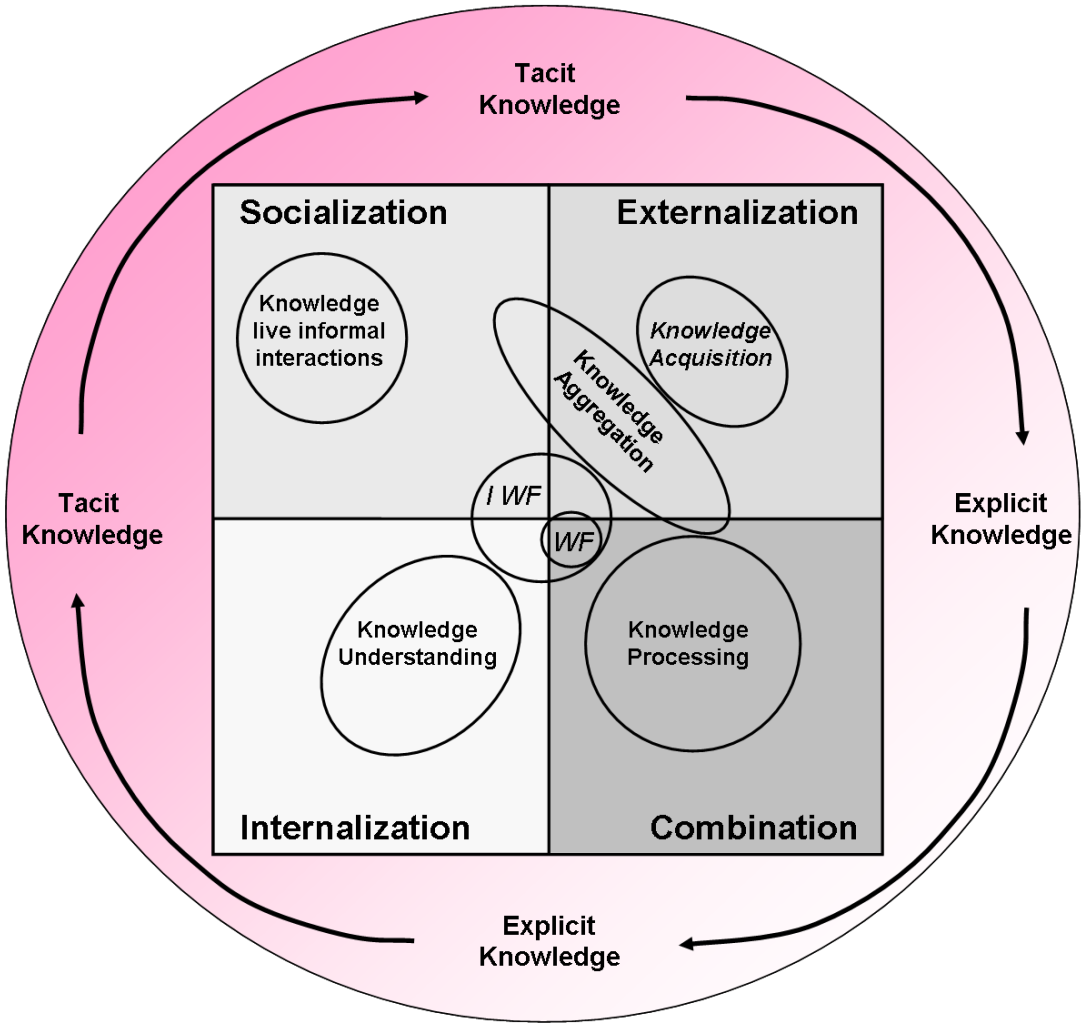

Figure 2.7: Workflow and SECI model.

2.3. Social Networks and Best Practice Networks

The Social Networks and the Best Practice Networks include many of the features that have been previously described. These complex systems are web platforms often realized customizing some of the most diffused Content Management Systems (CMS) or Learning Content Management Systems (LCMS). Many widespread Social Networks (SN) are mainly focused on contents (e.g., [YouTube], [Flickr], [LastFm]) [KraNac], whereas others are more focused on establishing relationships among users ([Facebook], [MySpace], [Orkut], [Friendster]), [NaRaCh]. SN analysis is typically focused on analyzing relationships among users, users and groups, in order to identify which are the most central users and groups, and, on the other hand, which are those that are frozen out, namely being those running the risk of losing interest in the network activities due to a serious lack of involvement, [BCFNPPa]. On the other hand, Best Practice Networks, BPN, are thematic social networks where smaller groups of users share content, contacts and information for a common goal. Some of them are set up by thematic communities, associations and large companies to capitalize skills and knowledge, by facilitating communications and interactions among personnel and the growth of internal content repository and knowledge.

An analysis related to social networking services provided by the Social Networks (both oriented on Users and on Content) and by the Best Practice Networks, and related to the Tacit and Implicit Knowledge involved in the services has been done. The comparison includes the list of functionalities related to users and social interaction and functionalities related to content fruition.

In the following sections, the relation among the actions that the users can do on the platforms (that obviously depend on the services offered by the platforms itself) and the kind of knowledge transformations involved during these actions is described. Both the Social Networks and the BPNs provide to the users a set of services, allowing them to make actions on the portal. Each *Action* has its flow of knowledge: starts using the platform services and ends creating new knowledge.

To better understand this analysis, it is useful to make some examples and take into account the Legend here after.

Action: Write a public Web Page.

- **Start**: the user starts from its own knowledge (tacit), uses editing tools and creates a web page on the platform (explicit). So the passage of knowledge is from tacit to explicit: T→E;
- **Result**: this is the result of the user action. The user has produced a web page that can be visible to all the users of the platform (explicit); the users can read it and get information (tacit): E→T.

Action: Recommendations

- **Start**: the platform produces and offers to the user a list of recommendations (explicit) and the user sees them. In this case the knowledge flow is from explicit to explicit: E→E;

- **Result**: the user reads the recommendations (explicit), acquires and reworks the information provided by the platform (tacit): E→T.

Legend	
O	Object, content
U	User
Y	Yes, the service is present
N	No the service is missing
A	Audio
V	Video
D	Document
I	Image
M	Multimedia
T	Tacit knowledge
E	Explicit knowledge
UGC	User Generated Content

Moreover in Figure 2.8 are depicted the four different kind of knowledge transformation that are taken into account to describe the services offered to the users from the Social Networks and the Best practice Networks:

- 1) from T→E to E →T: users make actions on the portal (uploading documents, making comments on content, etc.);
- 2) from T→E to E→E: also in this case users make actions on the portal (manage IPR on content, write lessons, etc.);
- 3) from E→E to E→E: the portal make internal actions on the portal (managing content, workflow support, etc.);
- 4) from E→E to E→T: portal produces stimulus to the users (making recommendations, or list of users, etc.).

It has to be noticed that:

- All the arrows (type 1, 4) that end in the Internalization quadrant can be interpreted as stimulus provided to the users by other users (type 1) or by the platform (type 4);
- All the arrows that start from the Externalization quadrant can be interpreted as actions made by the users on th platform (type 1,2);
- No arrows are present in the Socialization quadrant (from tatcit to tacit, T→T). This because it involves human live interactions not mediated from the ICTs. However It can be supposed, but it can not be measured, that a SN or a BPN gives stimulus to the users also for this kind of knowledge flow.

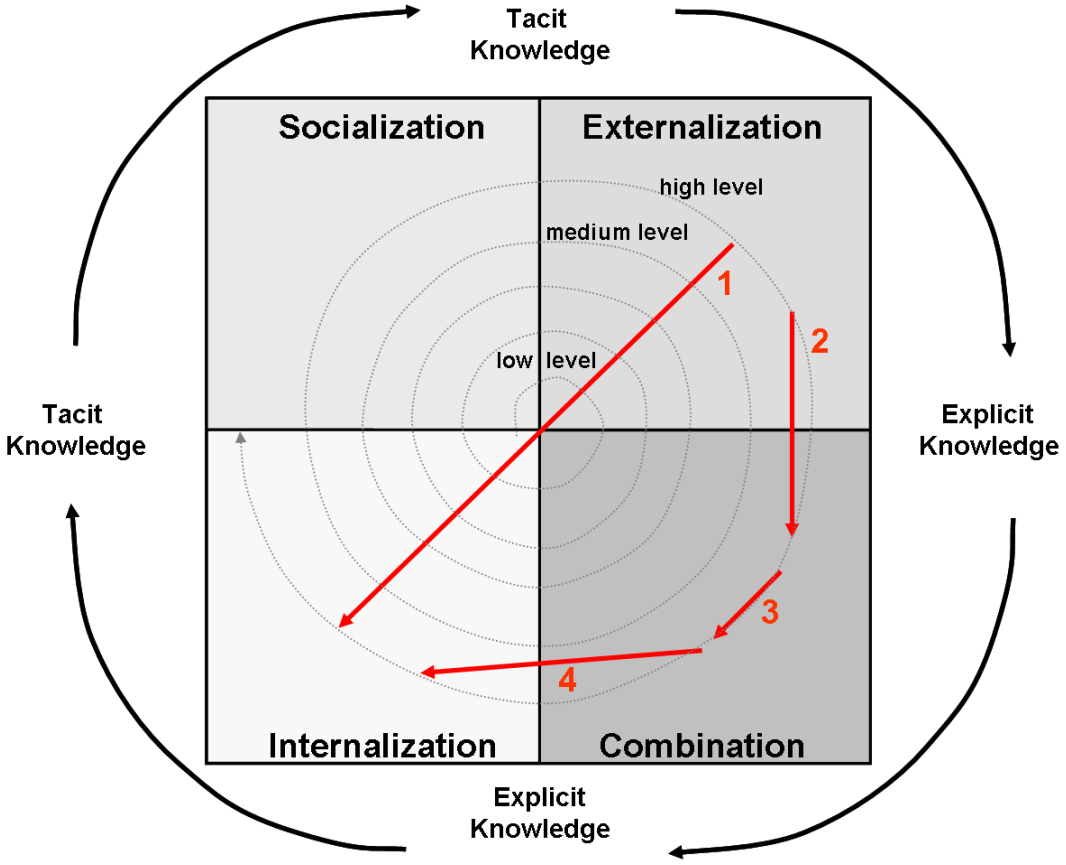

Figure 2.8: kinds of knowledge transformation.

2.3.1. Social Networks oriented to content

An analysis related to social networking services provided by Social Networks oriented to content and related to the tacit and implicit knowledge involved in the services has been done. The comparison includes the list of functionalities related to users and social interaction (see Table 2.7) and the functionalities related to content fruition (Table 2.8).

Social Networks oriented to Content: Content Services						
	Start			Result		Importance level (high, medium, low)
	Kind of user (Individual, Group, Portal)	Action		Result of the action	Kind of user (Individual, Group of users, Portal)	
Multimedia, cross-media UGC (upload)	Individual	E →E		E →T	All users	high
Audio, Video, Images, Doc	Individual	T→E		E →T	All users	high

(update)						
Moderated UGC	Individual	E →E		E →E	Group of users	high
Query on content	Individual	T→E		E →T	Individual	high
Comments on Content	Individual	T→E		E →T	All users	low
Ranking and voting	Individual	T→E		E →T	All users	low
General Recommendation O	Platform	E →E		E →T	All users	medium
Recommendation O→U	Platform	E →E		E →T	All users	medium
Recommendation O→O	Platform	E →E		E →T	All users	high
Taxonomy for content	Platform	E →E		E →T	All users	high
Play Lists of content	Individual	T→E		E →T	All users	high
RSS Feeds for content	Platform	E →E		E →T	All users	medium
Links with other SN	Individual	T→E		E →E	All users	low
Mobile Support	Platform	E →E		E →T	All users	medium
DRM/CAS Support, IPR mng	Individual	T→E		E →E	All users, group of users (it depends on privileges/licences, etc.)	high
Geo Tagging and QR	Individual	T→E		E →E	All users	low
e-learning courses (e.g. teacher writes a lesson)	Individual	T→E		E →E	Group of users (students)	medium (it depends on the context)
Workflow	Platform	E →E		E →E	All users	high

Table 2.7: Social Networks oriented to Content: Content Services.

Social Networks oriented to Content: User Services						
	Start			Result		Importance level (high, medium, low)
	Kind of user (Individual, Group, Portal)	Action : Production (creation, write, etc)		Action: Usage (read, play, access, etc.)	Kind of user (Individual, Group, Portal)	
User profile, descriptors	Individual	T→E		E →E	All users/ group of users/ Platform (it depends on information type)	medium
Friends/ colleagues	Individual	T→E		E →E	Individual (friend/ colleague)	low
Query on Users	Individual	T→E		E →E	Individual	low
Groups of users, Forums, etc.	Individual	T→E		E →T	Group of users	low
Multilingual pages (write)	Individual	T→E		E →T	All users	low
Invitations of users	Individual	T→E		E →E	Individual (invited user)	low
Chats, on line, messages, e-mail, etc.	Individual	T/E→E		E →E	Individual	low
Recommendation U→U	Platform	E →E		E →T	All users	low
User Lists, gen rec. of users	Platform	E →E		E →T	All users	low
Taxonomy on Users	Platform	E →E		E →T	All users	low
Events	Individual	T→E		E →T	Group of users	low
Group and user statistics	Platform	E →E		E →T	All users/ group of users/ Individual (it	low

					depends on the statistic type)	
Suggestions	Individual	T→E		E →E	Individual	low
Intelligent workflow	Platform/ group of users (those involved in the workflow)	E/T →E		E →E	All users	medium

Table 2.8: Social Networks oriented to Content: User Services.

In Figure 2.9, the arrows that represent the kinds of knowledge transformation involved in the services provided to the users in the Social Network oriented to Content are depicted. It can be seen that most of the arrows with the high and medium value:

- end on the Internalization quadrant;
- come from the Combination quadrant (or are entirely into it).

This reflects the fact that the Social Networks oriented to the Content have the aim to work on content and produce stimulus to the users such as recommendations on the content managed, etc.

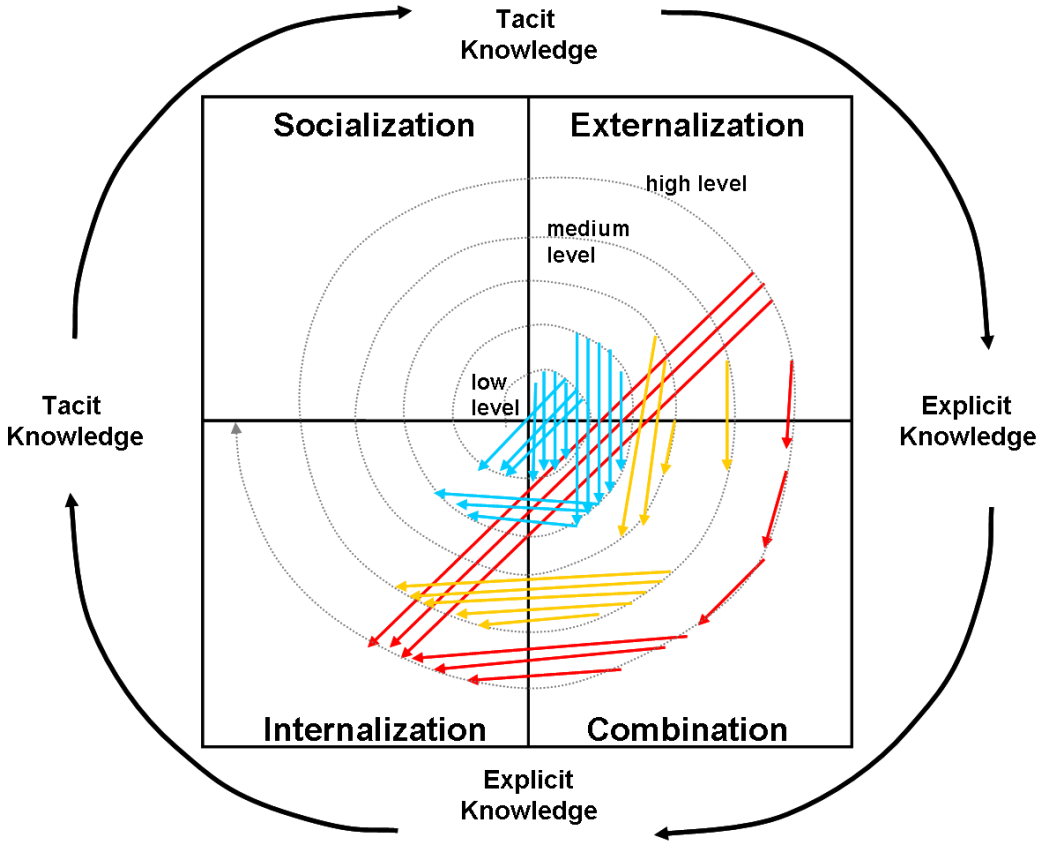

Figure 2.9: Social Networks oriented to Content and SECI model.

YouTube

Figure 2.9, represents a generic Social Network oriented to the content: is an optimistic evaluation. It is clear, that not all the Social Networks provide all the services that have been analysed. For this reason, it can be really useful to proceed considering one of the most diffused Social Networks oriented to the content: YouTube.

YouTube: Content Services						
	Start			Result		Importance level (high, medium, low)
	Kind of user (Individual, Group, Portal)	Action		Result of the action	Kind of user (Individual, Group of users, Portal)	
Multimedia, cross-media UGC (upload)	Individual	E →E		E →T	All users	high
Audio, Video, Images, Doc (update)	Individual	T→E		E →T	All users	high (only video)
Moderated UGC	Individual	E →E		E →E	Group of users	high
Query on content	Individual	T→E		E →T	Individual	high
Comments on Content	Individual	T→E		E →T	All users	High
Ranking and voting	Individual	T→E		E →T	All users	Medium
General Recommendation O	Platform	E →E		E →T	All users	High
Recommendation O→U	Platform	E →E		E →T	All users	High
Recommendation O→O	Platform	E →E		E →T	All users	High
Play Lists of content	Individual	T→E		E →T	All users	High
RSS Feeds for content	Platform	E →E		E →T	All users	Medium
Links with other SN	Individual	T→E		E →E	All users	Low

Mobile Support	Platform	E →E		E →T	All users	High
DRM/CAS Support, IPR mng	Individual	T→E		E →E	All users, group of users (it depends on content)	medium
Workflow	Platform	E →E		E →E	All users	High

Table 2.9: YouTube Content Services.

YouTube: User Services						
	Start			Result		Importance level (high, medium, low)
	Kind of user (Individual, Group, Portal)	Action : Production (creation, write, etc)		Action: Usage (read, play, access, etc.)	Kind of user (Individual, Group, Portal)	
User profile, descriptors	Individual	T→E		E →E	All users	low
Friends/ colleagues	Individual	T→E		E →E	Individual (friend)	low
Groups of users, Forums, etc.	Individual	T→E		E →T	Group of users	low
Multilingual pages (write)	Individual	T→E		E →T	All users	low
Invitations of users	Individual	T→E		E →E	Individual (invited user)	low
Content statistics	Platform	E →E		E →T	All users	high
Suggestions	Individual	T→E		E →E	Individual	low

Table 2.10: YouTube User Services

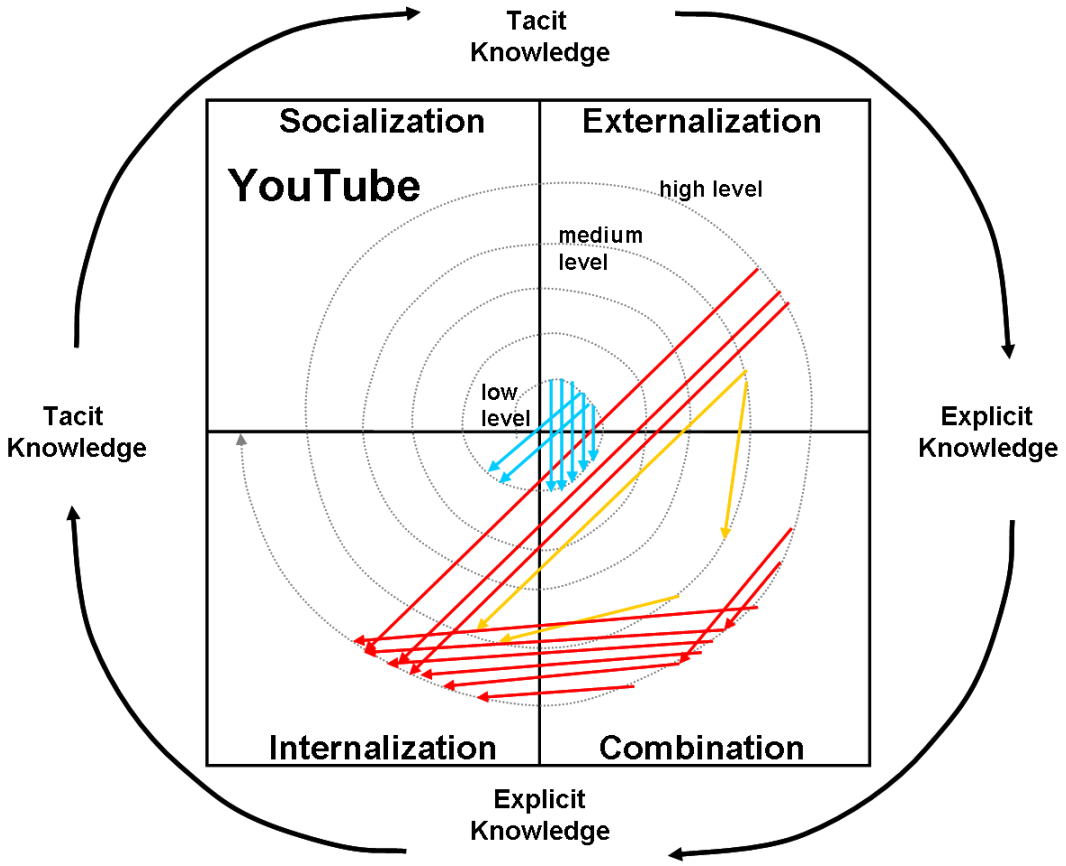

Figure 2.10: YouTube and SECI model.

Observing Figure 2.10, it appears clear that YuTube is a platform oriented to give personal stimulus to the users. In fact people usually goes on YouTube with the aim to see a video clip, to listen to a song, to find video associated to others, to learn new knowledge reading manuals or viewing instructional video, etc.

2.3.2. Social Networks oriented to users

An analysis related to social networking services provided by Social Networks oriented to users and related to the tacit and implicit Knowledge involved in the services has been done. The comparison includes the list of functionalities related to users and social interaction (see Table 2.11) and functionalities related to content fruition (see Table 2.12).

Social Networks oriented to Users: Content Services						
	Start			Result		Importance level (high, medium, low)
	Kind of user (Individual, Group, Portal)	Action		Result of the action	Kind of user (Individual, Group of users, Portal)	
Multimedia, cross-media UGC	Individual	E →E		E →T	All users	medium

(upload)						
Audio, Video, Images, Doc (update)	Individual	T→E		E →T	All users	medium
Moderated UGC	Individual	E →E		E →E	Group of users	medium
Query on content	Individual	T→E		E →T	Individual	medium
Comments on Content	Individual	T→E		E →T	All users	high
Ranking and voting	Individual	T→E		E →T	All users	high
General Recommendation O	Platform	E →E		E →T	All users	low
Recommendation O→U	Platform	E →E		E →T	All users	low/ medium
Recommendation O→O	Platform	E →E		E →T	All users	low
Taxonomy for content	Platform	E →E		E →T	All users	low/ medium
Play Lists of content	Individual	T→E		E →T	All users	medium/ high
RSS Feeds for content	Platform	E →E		E →T	All users	high
Links with other SN	Individual	T→E		E →E	All users	high
Mobile Support	Platform	E →E		E →T	All users	high
DRM/CAS Support, IPR mng	Individual	T→E		E →E	All users, group of users (it depends on privileges/licences, etc.)	medium
Geo Tagging and QR	Individual	T→E		E →E	All users	high
e-learning courses (e.g. teacher writes a lesson)	Individual	T→E		E →E	Group of users (students)	medium (it depends on the context)
Workflow	Platform	E →E		E →E	All users	medium

Table 2.11: Social Networks oriented to Users: Content Services.

Social Networks oriented to Users: User Services						
	Start			Result		Importance level (high, medium, low)
	Kind of user (Individual, Group, Portal)	Action : Production (creation, write, etc)		Action: Usage (read, play, access, etc.)	Kind of user (Individual, Group, Portal)	
User profile, descriptors	Individual	T→E		E →E	All users/ group of users/ Platform (it depends on information type)	high
Friends/ colleagues	Individual	T→E		E →E	Individual (friend/ colleague)	high
Query on Users	Individual	T→E		E →E	Individual	high
Groups of users, Forums, etc.	Individual	T→E		E →T	Group of users	high
Multilingual pages (write)	Individual	T→E		E →T	All users	high
Invitations of users	Individual	T→E		E →E	Individual (invited user)	high
Chats, on line, messages, e-mail, etc.	Individual	T/E→E		E →E	Individual	high
Recommendation U→U	Platform	E →E		E →T	All users	high
User Lists	Platform	E →E		E →T	All users	high
Taxonomy on Users	Platform	E →E		E →T	All users	high
Events	Individual	T→E		E →T	Group of users	high
Group and user statistics	Platform	E →E		E →T	All users/ group of users/ Individual (it depends on the statistic type)	medium
Suggestions	Individual	T→E		E →E	Individual	high
Intelligent	Platform/	E/T →E		E →E	All users	high/me

workflow	group of users (those involved in the workflow)					dium

Table 2.12: Social Networks oriented to Users: User Services.

In Figure 2.11, the arrows that represent the kinds of knowledge transformation involved in the services provided to the users in the Social Network oriented to Users are depicted. It can be seen that most of the arrows start from the Externalization quadrant, this reflect the fact that the Social Network oriented to the Users have the aim to involve the people on the portal.

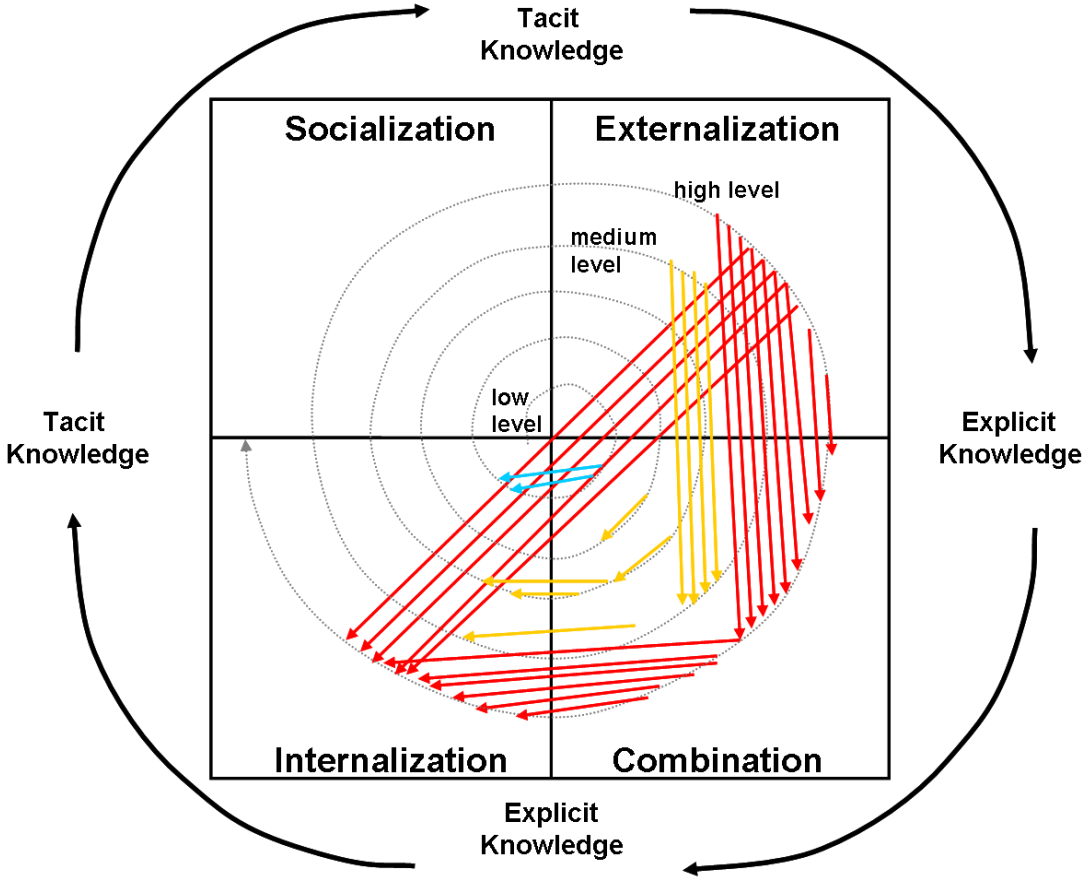

Figure 2.11: Social Network oriented to users and SECI model..

Facebook

Figure 2.11, represents a generic Social Network oriented to the user: is an optimistic evaluation. It is clear, that not all the Social Networks provide all the services that have been analysed. For this reason, it can be really useful to proceed considering one of the most diffused Social Networks oriented to the user: Facebook.

Facebook: Content Services						
	Start			Result		Importance level (high, medium, low)
	Kind of user (Individual, Group, Portal)	Action		Result of the action	Kind of user (Individual, Group of users, Portal)	
Multimedia, cross-media UGC (upload)	Individual	E →E		E →T	All users	Medium (only multimedia)
Audio, Video, Images, Doc (update)	Individual	T→E		E →T	All users	medium (only images, documents and video)
Comments on Content	Individual	T→E		E →T	All users	high
Ranking and voting	Individual	T→E		E →T	All users	medium
General Recommendation O	Platform	E →E		E →T	All users	low
Recommendation O→U	Platform	E →E		E →T	All users	medium
Links with other SN	Individual	T→E		E →E	All users	medium
Mobile Support	Platform	E →E		E →T	All users	high
Workflow	Platform	E →E		E →E	All users	medium

Table 2.13: Facebook: Content Services.

Facebook: User Services						
	Start			Result		Importance level (high, medium, low)
	Kind of user (Individual, Group, Portal)	Action : Production (creation, write, etc)		Action: Usage (read, play, access, etc.)	Kind of user (Individual, Group, Portal)	
User profile, descriptors	Individual	T→E		E →E	All users/ group of users (it depends on the users'	high

					choices)	
Friends/ colleagues	Individual	T→E		E →E	Individual (friend)	high
Query on Users	Individual	T→E		E →E	Individual	high
Groups of users, Forums, etc.	Individual	T→E		E →T	Group of users	high
Multilingual pages (write)	Individual	T→E		E →T	All users	medium
Invitations of users	Individual	T→E		E →E	Individual (invited user)	high
Chats, on line, messages, e-mail, etc.	Individual	T/E→E		E →E	Individual	high
Recommendation U→U	Platform	E →E		E →T	All users	high
User Lists	Platform	E →E		E →T	All users	high
Events	Individual	T→E		E →T	Group of users	high
Suggestions	Individual	T→E		E →E	Individual	high

Table 2.14: Facebook: Users Services.

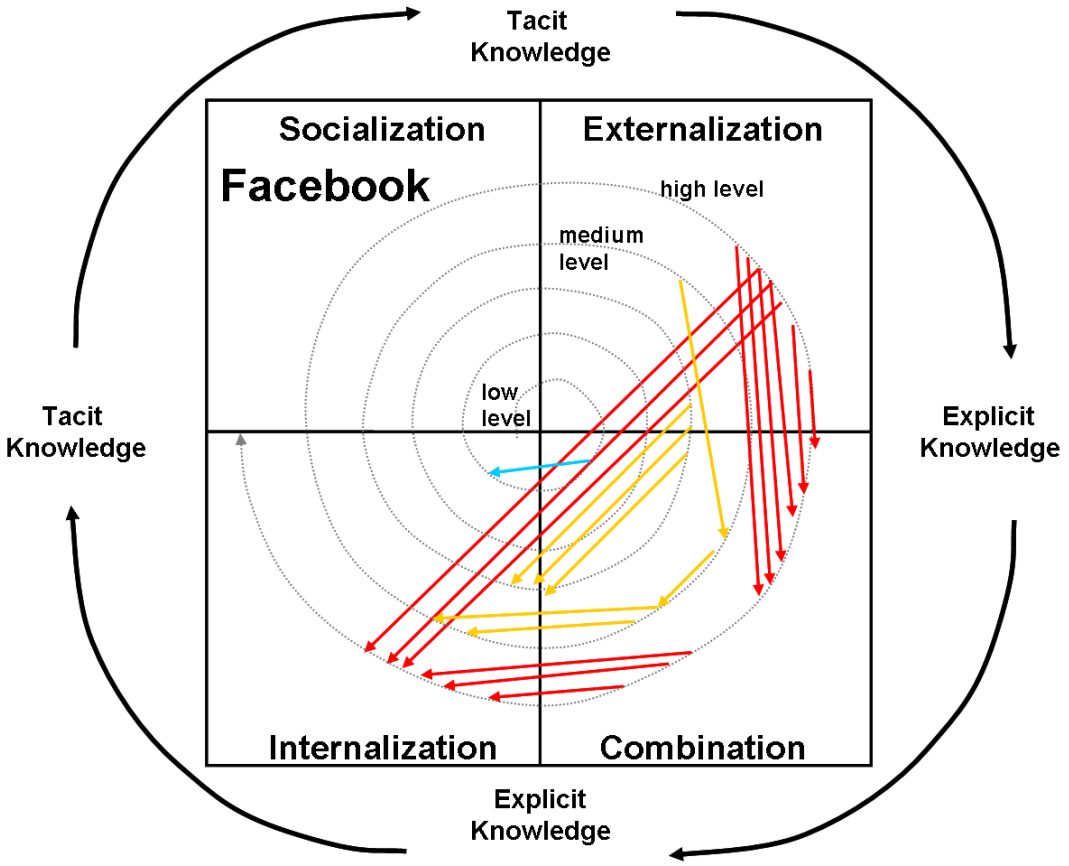

Figure 2.12: Facebook and SECI model.

Observing Figure 2.12, it appears clear that Facebook is a platform oriented to stimulate the socialisation (mediated by the ICTs) of the people among them. In fact people goes on Facebook to create its group of friend, to meet new people and have information on them, to chat and communicatwith its friends, etc. But it is also relaevant to note that users also pose new knowledge on the portal, with the aim of socialise: creating their personal wall, uploading photo and video, making comments on the content of the friends, etc.

2.3.3. Best Practice Networks

It is relevant to summarize the main differences of Best practice Networks, BPN, with respect to classical massively widespread SNs:

- BPNs present a larger percentage of active contributors, for both content and comments with respect to SN, the size of the internal communities are smaller. Users of BPN are typically more motivated to participate since BPN are strongly thematic and connected to the user's work. In SNs, the effective/active contributors are typically much less than 10%, all the others are merely observers;
- BPNs are typically without advertising, since the business model is designed to provide a service, in order to understand user needs and it is not meant to make earnings with product advertising on a larger number of users;

- BPNs have to cope with a large variety of content types and not only with videos and images like it occurs on large SNs. Most SNs can manage only simple audiovisual content (e.g., images or videos, which are the simplest content type to be generated by users, for example via digital camera and/or phones). Many BPNs have to be able to cope with complex interactive content for education, edutainment, and/or entertainment experiences. This means having more complex semantic descriptors for those content items. At present, there are many multimedia and cross content formats;
- BPNs have to cope with a large variety of semantic computing activities/algorithms: for content adaptation and processing (also due to the complexity of content) of user contributions: user generated content, indexing of user contributions/comments, etc. In BPN, the number of users and items is smaller, so that the complexity of semantic processing in terms of triples and ontology complexity may be larger, though still economically viable;
- BPNs can typically cover more platforms for providing services and content distribution; for example: PC, windows mobile PDAs, i-Phone, etc.;
- BPNs have more complex content and services, therefore users/content profiles/descriptors have to consider static as well as dynamic aspects [BeBFNP]. In BPN, like in SN, dynamic aspects are much more relevant since they may be continuously updated.

Best Practice Networks: Content Services						
	Start			Result		Importance level (high, medium, low)
	Kind of user (Individual, Group, Portal)	Action		Result of the action	Kind of user (Individual, Group of users, Portal)	
Multimedia, cross-media UGC (upload)	Individual	E →E		E →T	All users	high
Audio, Video, Images, Doc (update)	Individual	T→E		E →T	All users	high
Moderated UGC	Individual	E →E		E →E	Group of users	high
Query on content	Individual	T→E		E →T	Individual	high
Comments on Content	Individual	T→E		E →T	All users	high
Ranking and voting	Individual	T→E		E →T	All users	high

General Recommendation O	Platform	E →E		E →T	All users	high
Recommendation O→U	Platform	E →E		E →T	All users	high
Recommendation O→O	Platform	E →E		E →T	All users	high
Taxonomy for content	Platform	E →E		E →T	All users	high
Play Lists of content	Individual	T→E		E →T	All users	high
RSS Feeds for content	Platform	E →E		E →T	All users	high
Links with other SN	Individual	T→E		E →E	All users	high
Mobile Support	Platform	E →E		E →T	All users	high
DRM/CAS Support, IPR mng	Individual	T→E		E →E	All users, group of users (it depends on privileges/licences, etc.)	high
Geo Tagging and QR	Individual	T→E		E →E	All users	high
e-learning courses (e.g. teacher writes a lesson)	Individual	T→E		E →E	Group of users (students)	medium (it depends on the context)
Workflow	Platform	E →E		E →E	All users	high

Table 2.15: Best Practice Networks: Content Services.

Best Practice Networks: User Services						
	Start			Result		Importance level (high, medium, low)
	Kind of user (Individual, Group, Portal)	Action : Production (creation, write, etc)		Action: Usage (read, play, access, etc.)	Kind of user (Individual, Group, Portal)	
User profile, descriptors	Individual	T→E		E →E	All users/ group of users/ Platform (it depends on information	high

					type)	
Friends/ colleagues	Individual	T→E		E →E	Individual (friend/ colleague)	high
Query on Users	Individual	T→E		E →E	Individual	high
Groups of users, Forums, etc.	Individual	T→E		E →T	Group of users	high
Multilingual pages (write)	Individual	T→E		E →T	All users	high
Invitations of users	Individual	T→E		E →E	Individual (invited user)	high
Chats, on line, messages, e-mail, etc.	Individual	T/E→E		E →E	Individual	high
Recommendation U→U	Platform	E →E		E →T	All users	high
User Lists	Platform	E →E		E →T	All users	high
Taxonomy on Users	Platform	E →E		E →T	All users	high
Events	Individual	T→E		E →T	Group of users	high
Group and user statistics	Platform	E →E		E →T	All users/ group of users/ Individual (it depends on the statistic type)	high
Suggestions	Individual	T→E		E →E	Individual	high
Intelligent workflow	Platform/ group of users (those involved in the workflow)	E/T →E		E →E	All users	medium/ high

Table 2.16: Best Practice Networks: User Services.

In Figure 2.13**Figure 2.11**, the arrows that represent the kinds of knowledge transformation involved in the services provided to the users in a Best Practice Network are depicted.
It can be seen that the arrows:

- have in prevalence an high value;
- their presence in the quadrants is balanced: only the Socialization quadrant is excluded. This because it involves human interactions not mediated from the ICTs.

A BPN provide many services to the users and stimulate them in a constructively manner: in fact, the users make a large number of actions on these kind of portals.

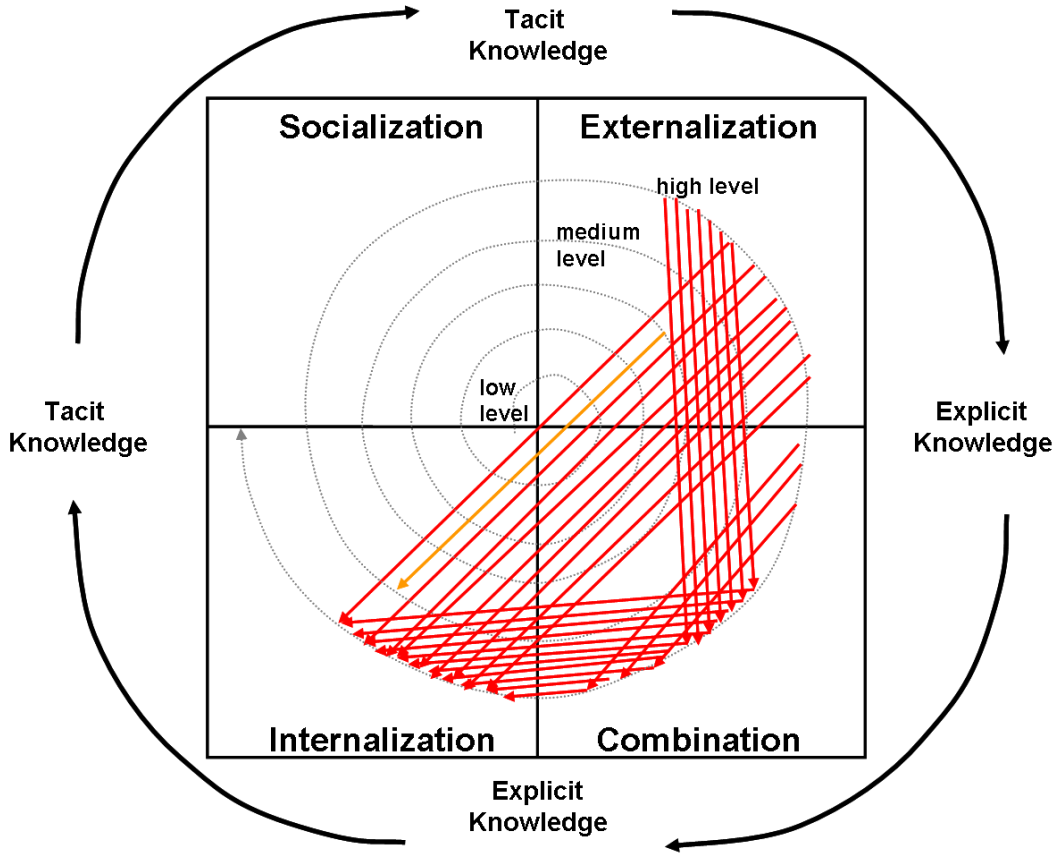

Figure 2.13: Best Practice Networks and SECI model.

2.4. Comparison of the ICTs

In the previous paragraphs technological tools and frameworks, that are now more common for the dissemination, management and development of knowledge, have been described. These instruments have been compared with the SECI model in

order to understand to which of the four modes of knowledge conversion they are able to cover.

After the analysis made on the Social Networks and on the Best Pactice Networks is possible to represent how these kind of platforms (see Figure 2.14) cover the four quadrants of the SECI model.

The SNs oriented to content have the aim to work on content and produce stimulus to the users: these plaforms maily cover the Combination and Internalization quadrants.
The SNs oriented to users have the aim to involve the people on the portal to stimulate their socialisation (mediated by the ICTs). In order to do this they maily cover the quadrants involving the tacit knowledge that are Internalization and Externalization, but also in a minor percentage, the Combination quadrant.
The Best Practice Networks are more balanced respect to the SNs, in fact all the three quadrants involving the explicit knowledge are covered: Internalization, Externalization and Combination.
Moreover it has been supposed that a SN or a BPN gives stimulus to the users also for the tacit to tacit knowledge flow (T→T): these platforms stimulate the users to make actions that can be situated in the Socialization quadrant (in a low percentage respect to the other modalities of knowledge flow).

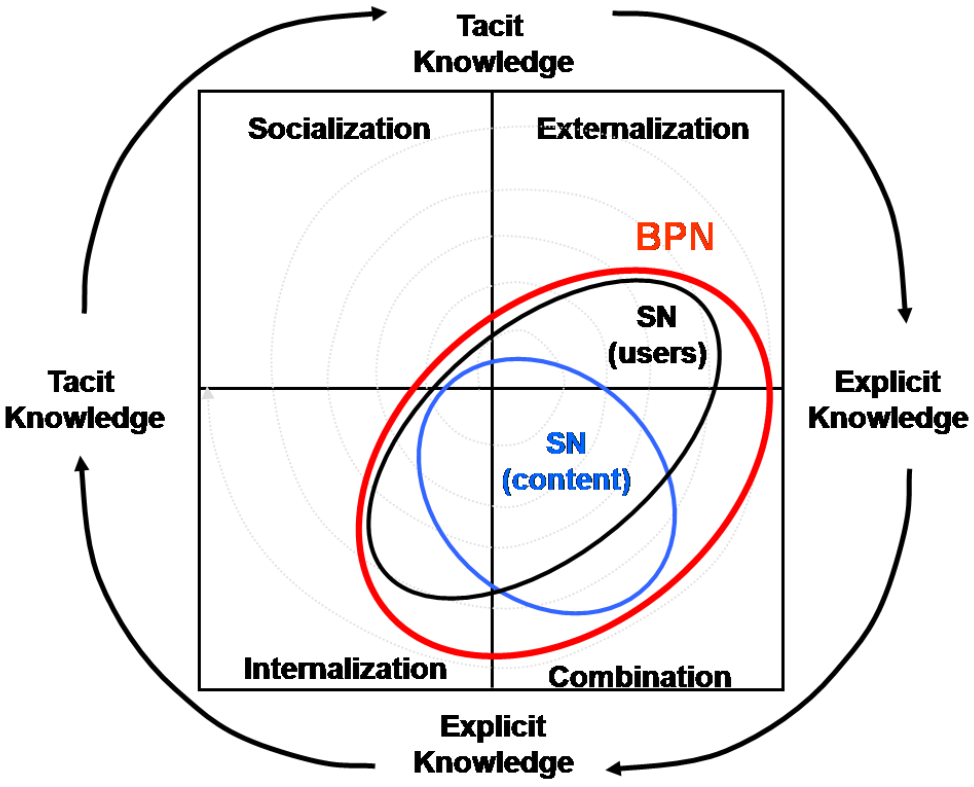

Figure 2.14: Social Networks, Best Practice Networks and SECI model.

After this analysis, the conclusion is that the best way to stimulate the development and creation of knowledge in an unique platform is to develop a Best Practice Network (see Figure 2.14). This is exactly what has be done: in the following chapter the architecture of the Best Practice Network developed is described in details.

3. Best Practice Network General Architecture

As described in the previous chapter, in order to create an infrastructure that could integrate all the features that are at the base of the SECI model and that takes into account the importance of the communities of practice in the knowledge creation and transfer processes, it was chosen to design and implement a Best Practice Network.
It starts describing the connections between the Best Practice Network and the theories, analysed in the earlier chapters, for the knowledge management; after that the real description of the architecture is done. The following sections are related to description of the Best Practice Network Architecture. As it will be clear after reading the architecture description, the BPN developed can be easily customised and applied to many different contexts: in order to better show its potentialities, in the next chapters different scenarios are described.

3.1. Best Practice Network developed and SECI model

As anticipated in the previous chapter, the technology that allows to embrace the fundamental characteristics for optimal management of the flow of both tacit and explicit knowledge within an organization and among different organizations is a Best Practice Network (BPN). A BPN provides the necessary tools to individuals to develop the knowledge in the best way, leveraging community of practice, using the individual know-how, increasing the personal skills by learning or through new perspectives, using the resources that the organizations and the people involved make them available.
As noticed in chapter 1, referring to the SECI model, there are four modes of knowledge conversion. They are: (1) socialisation (from tacit knowledge to tacit knowledge); (2) externalisation (from tacit knowledge to explicit knowledge); (3)

combination (from explicit knowledge to explicit knowledge); and (4) internalisation (from explicit knowledge to tacit knowledge), [NonToyKon].

The BPN developed, Figure 3.1, provides many services to the users and stimulate them in a constructively manner. As appears in the following figure, the users make a large number of actions on the portal and it provides to them a lot of services.
Although users could be further stimulated by the platform in order to intensify the production of knowledge.

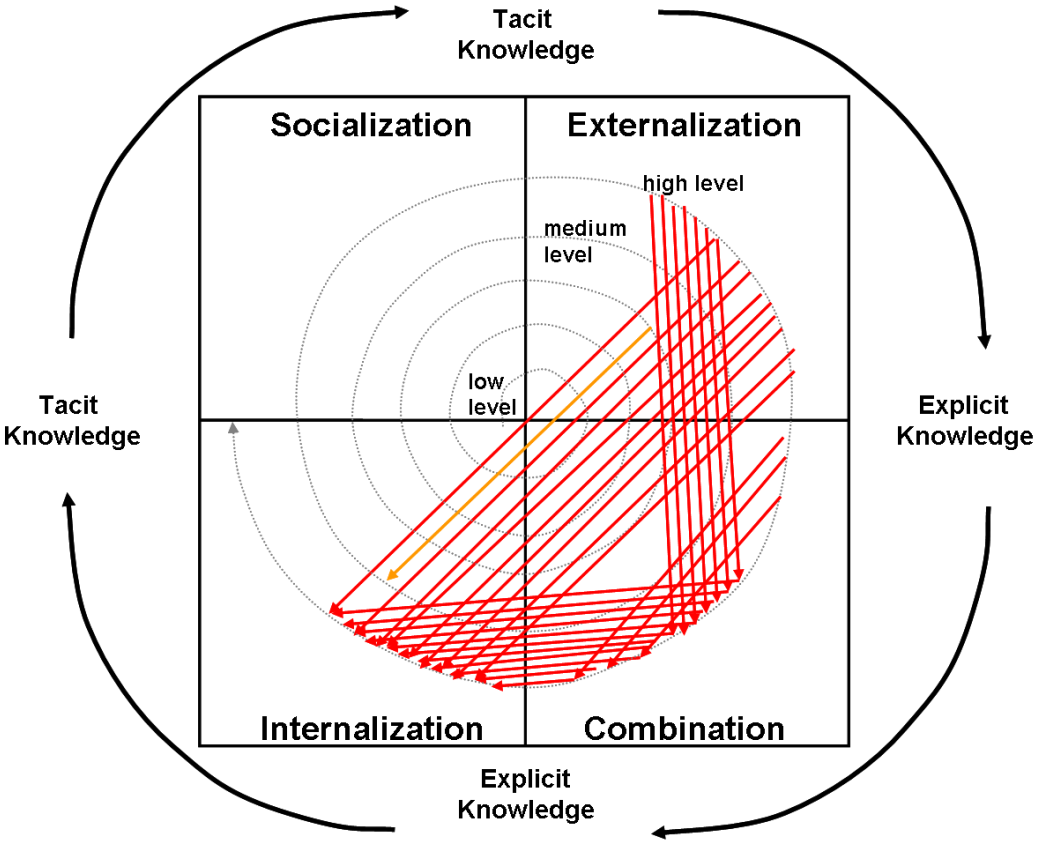

Figure 3.1: Developed Architecture and SECI model.

In Figure 3.1 it is represented only the Epistemological dimension in which is represented only the relation between the explicit and tacit knowledge; it is possible to view a graph, Figure 3.2 in which the Nonaka's spiral is represented taking into account both the epistemological dimension and the ontological one: the knowledge management into an organization is a continuous stream that includes an exchange of information among individuals (groups, organizations, inter-organization) and an internal reworking of the intellectual capital (personal).
According to this perspective, the BNP has been developed giving fundamental importance to individual initiative, the exchange of information among people and in groups, valorising and stimulating the most creative people offering them a creative context in which to create knowledge; using the best strategies and advanced technology to acquire, retrieve, extract, protect the multimedia contents available in the network.

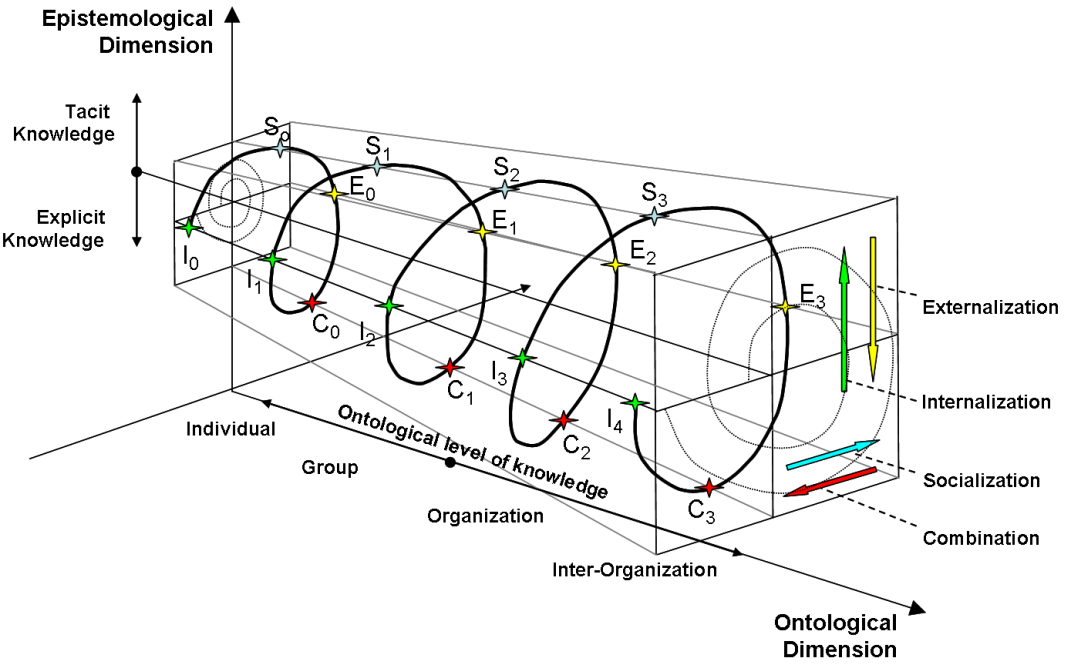

Figure 3.2: Epistemological and ontological dimension of the Nonaka's spiral.

Another aspect that has been taken into account during the BPN development is related to the knowledge assets continuously evolving. According to the Nonaka's definition, the knowledge assets are the resources and the methodologies adopted during the knowledge creation process. To understand how knowledge assets are created, acquired and exploited, they can be categorised into four types:

- Experiential, consists of the shared tacit knowledge that is built through shared hands-on experience amongst the members of the organisation, and between the members of other organisation (e.g. skills, know-how, emotional knowledge, physical knowledge);
- Conceptual, consists of explicit knowledge are articulated through images, symbols and language (e.g. brand equity, concepts or designs);
- Systemic, consists of systematised and packaged explicit knowledge, such as: specifications, manuals, etc. It is fundamental that All these intellectual properties have to be legally protected, for example using licences and patents (e.g. documents, specifications, manuals, etc, that have to be legally protected, for example using licences and patents);
- Routine, consists of the tacit knowledge that is routinised and embedded in the actions and practices of the organisation (e.g. know-how in daily operations, organisational routine, organizational culture).

The BPN developed covers all the aspects previously mentioned, moreover another important aspect for the development and dissemination of knowledge described in the SECI model is related to the Knowledge Assets (KA) that, according to the Nonaka's definition, are firm-specific resources that are indispensable to create values for the firm, [NonToyKon]. The most important assets of an organization develop and sustain its competitive advantage, moreover they are constantly

evolving To understand how knowledge assets are created, acquired and exploited, they can be categorised into four types: Experiential, Conceptual, Systemic and Routine (for more details see chapter 1). Conceptual and Systemic knowledge assets regard the explicit knowledge treatment while Experiential and Routine knowledge assets concern the assimilation and the exchange of tacit knowledge.
In Figure 3.3, the Knowledge Assets in an organization that makes use of a Best Practice Network to organize its work, are represented. It comprehends:

- the organization members (M) each of them with its tacit knowledge;
- the resources (or explicit knowledge)

The resources in all their forms (images, symbols, specifications, manuals, etc.) can be stored in database (Conceptual and Systemic KA) and interchanged among the members. The tacit knowledge can be shared, enriched or acquired through the consultation of the resources or using the services provided to the users by the BPN. This can happen both at individual level or also taking part on group discussions (Experiential and Routine KA). The groups can be composed by members of the same organizations or coming from different organizations.
In Fig. 3 it is also clear that the Knowledge Assets (KA) change during the time: it means that tacit and explicit knowledge are continuously evolving. For this reason is fundamental to use integrated strategies to take track of these changes (workflow management systems) has it happens in a BPN.

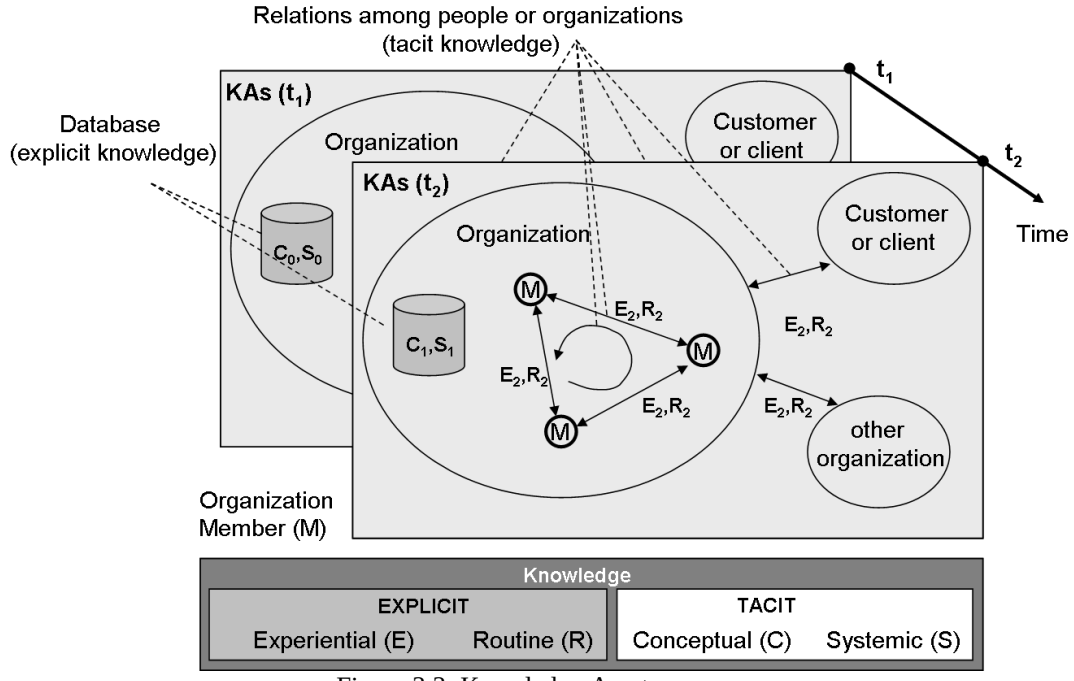

Figure 3.3: Knowledge Assets.

3.2. Best Practice Network Architecture

Best Practice Networks are thematic Social Networks focused on sharing common area works and goals. In a BPN smaller groups of users share content, contacts and information for a common goal. Some of them are set up by thematic communities, associations and large companies to capitalize skill and knowledge, by facilitating communications and interactions among personnel and the growth of internal content repository and knowledge. The analysis of user behaviour on social network is fundamental to tune services and stimulate the network growth. This is even more relevant for best practice collaborative networks where details about collaboration may lead to understand the effective activities and role of people in the several groups, [BCFNPPa].

The BPN Architecture model exploiting the cross media multi-channel distribution has been created to provide support in distributing cross media content towards different kinds of device: PC, PDA and Mobiles (smart phones, i-Phone, iPod, iPad). The simplified architecture of the BPN considered solution is depicted in Figure 3.4 where main modules are reported, [BeBFNP]:

- **User Management**: mainly for registration and profiling;
- **Mobile Support**: mobile content distribution and monitoring;
- **User Generated Content (UGC) Management**: is in connection with back office tools for content ingestion, processing and adaptation of the content;
- **Collaborative Support** among users: realised through groupware and communication and collaboration tools such as: social chat, grouping, messaging, notifications, etc.;
- **Indexing and Querying**: multilingual indexing, fuzzy support, semantic indexing, etc.
- **Recommendation support**: for content and users

All the capabilities are deployed and shown to the BPN users via different front-ends, specialized for device category: PC (Windows, Mac OS X, Linux, Unix, etc.), Smart phone, i-Phone with a dedicated Application (see Mobile Medicine i-Phone App [MobMedApp]), Windows Mobile PDA with a dedicated application called AxObjectFinder, [MobMedHelp].
The provided front-ends share the same database and user registration mechanism, and they may provide different services via different sets of languages. Moreover, different devices can support different modalities to access certain content: video may be accessed in download, progressive download (also called HTTP streaming), or even via P2P (only on XMF). These capabilities are only available for some devices and solutions. This is the typical complexity of a multi-channel, multi-device, content management and distribution system for a BPN.
In BPNs content types may range from single files: audio, video, images, documents and animations, simple slides, animations; to cross media such as: interactive guidelines/procedures, calculators and tools, questionnaires and data collection, leaning objects. The cross media content and file formats may be of

several kinds ranging over: MPEG-21[1], SCORM, XMF, HTML, SMIL, NewsML, SportML, METS, etc. ([SCORM], [XMF], [HTML], [SMIL], [METS]) In this context, cross media means objects produced as MPEG-21 files formats in the AXMEDIS[2] Project [BeBNR].

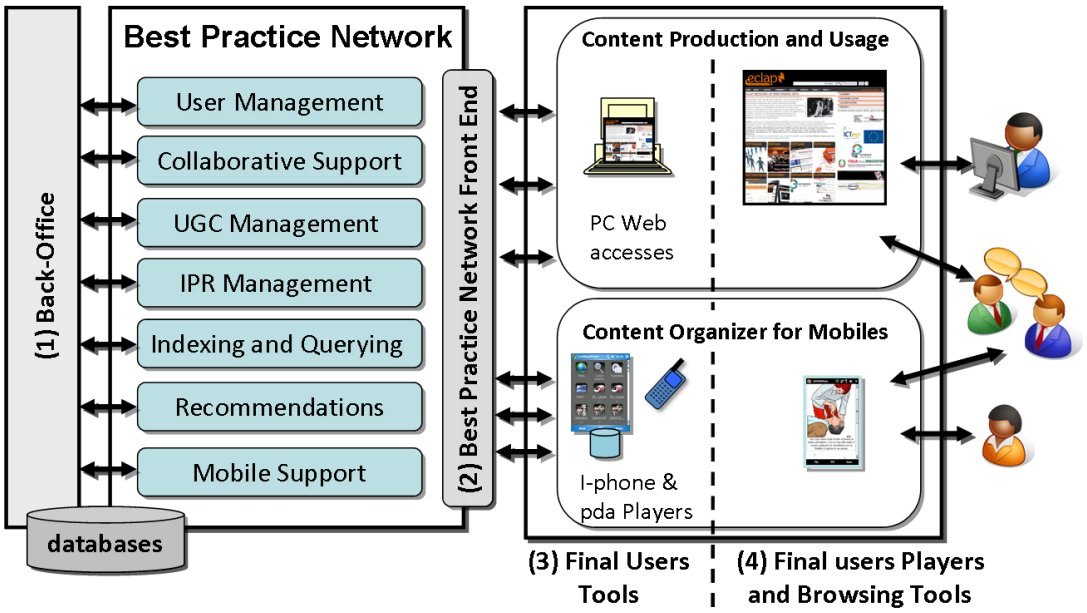

Figure 3.4: Simplified Best Practice Network Architecture.

Figure 3.4 reports the general simplified architecture of the Best Practice Network developed. It consists of a four layers architecture including:

- (1) a **Back-Office** for semantic computing;
- (2) a **Front-End** portal and services which performs a part of semantic computing activities and provides on-line services to the client side layers;
- (3) the **Final Users Tools** (Content Organizer)) for content management on mobile devices and content production on PC;
- (4) the **Final users Players and Browsing Tools** located on PC, PDA, i-Phone devices. In the case of PDA and i-Phone mobile devices, they are typically activated by the Content Organizer, while on PC, they are activated by the browser.

[1] MPEG-21 is focused on the standardization of the content description related to Digital Rights Management, DRM, aspects [LHJYPR], [NeRoV].

[2] AXMEDIS. AXMEDIS, [Axmedis] is a content media framework developed by an European Commission IST IP Research project. AXMEDIS extended MPEG-21 proposing content packaging, integrating presentation aspects, semantic computing, intelligence and behavioural capabilities [BeBNS].

3.2.1. Content Management System (CMS): Drupal

The Best Practice Network developed is based on the CMS Drupal [Drupal]. Drupal has been chosen (among other CMS) as starting point to realise the Best Practice Network basing on its features and taking into account also the following aspects:

- it is distributed under the General Public License (GPL)
- the languages used are: Php, J2EE, html, css, etc.)
- the main web server Installable on can be: Apache httpd, light or IIS
- it interfaces with Mysql or PostgreSql
- it works with different operating systems: Windows, Mac OSX, Linux
- it has an extended Community, that is composed of both developers and users
- it is modular: is composed of a set of default modules (each module implements one functionality or a group of functionalities on the CMS), the so called "Drupal Core Modules". Moreover Additional modules are developed and released in order to implement more functionalities.
- the Drupal core modules already provide the major features typical of a Best Practice Network:
    - users management (extended User Profiling, etc.);
    - content management (Knowledge acquisition, retrieval, extraction, protection, etc.);
    - social Interactions (groupware and communication and collaboration tools, e-learning, etc.);
    - multi-language;
- it is easy to customise it: creating new modules, interfacing it with other CMS, etc.

3.2.2. Best Practice Network Module and Services

The Best Practice Network manages WEB portal/pages for multi-channel distribution. A number of services are provided via web interface front-end of the BPN, while they are enabled by the back office. The **Best Practice Network Front-end** is the web-based multi-language interface providing the collaborative environment user interface and the content access according to several protocols: http, http download, progressive download for audio visual, and P2P towards PC and mobile devices. It recognizes PC and mobile device distributing the right content format and web pages in real time, [BeBFNP].
Thus, the most important services and modules are described in details:

- User Management module: allows the registration of new users , manages the user authentication, collects data on users: both static (user profile) and dynamic (user activities on the portal). Each user may:
    - establish connection with other users;

- be registered to multiple working groups on the BPN;
- provide/play content, make queries/comments/votes on content, etc.;
- send messages to colleagues or friends, receive notifications related to portal changes, etc.

The users may also to accept working with protected content according to IPR management models using the tools described in the IPR Management module.

- **Collaborative Support**: this module manages the BPN activities such as: web pages production, discussion forums and blogs, content enrichment, messaging, news, multilingual translations, etc. Most of these activities can be performed on both on PC and mobile devices.
- **UGC Management (User Generated Content)**: manages workflow activities of content upload, enrichment, review, acceptance, publication. Once the content is uploaded, it may be inspected or directly passed to the back office for the automated transcoding and formatting for the publication towards final users devices. During the content upload, a set of metadata, the group assignment of content, and the taxonomical classifications, are requested. A multilingual enrichment is also performed, assigning/estimating multiple languages for metadata, groups and taxonomy.
- **IPR Management (Intellectual Property Rights):** allows defining rules to access and exploit content on the portal and on mobiles devices. Two aspects have been developed:
  - **IPR rules**. This method derives from studies developed during the Axmedis Project and provides rules formalized in MPEG-REL (Rights Expression Language) and posted on the AXMEDIS DRM licensing server. The IPR module is based on AXMEDIS MPEG-21 DRM server ([Axmedis]) to manage content with control and protection. The AXMEDIS DRM may be used to enforce protection and release licenses at registered users with multiple rights and conditions according to MPEG-21 REL standard. It can be used as a simple CAS (Conditional Access Systems) or as a full DRM and permits to manage black lists of users, licenses, devices, etc. [LHJYPR], [NeRoVa] to revoke content access and/or rights. The content is protected via encryption, the registered users may use content only on certified players and if a suitable license is accessible on the AXMEDIS DRM licensing server. The AXMEDIS DRM includes: the licensing server, the verification and authentication server, and the certification authority.
  - **Licences**. A system to associate a licence to each content basing on: the access permissions that the user that have uploaded it on the Portal; the Rights that the user have on the digital content. This aspect is quite complex and is described in details in the ECLAP scenario [Eclap] (see chapter 5).
- **Indexing and Querying** data Collection. This module has the responsibility of collecting data for indexing the several kinds of content managed on the portal: content object for pc and mobile, web pages, forum topics, comments on any kind of content, etc. The queries can be

performed on the portal by using the above-mentioned modalities as simple text queries or as advanced queries specifying the modalities. A complete substring match and advanced detailed search are provided for specialists. The indexing is performed by using a fuzzy indexing for each language and a global index with a Lucene based solution [PiTa], [Lucene]. The massive activity of indexing and re-indexing are performed in the back office by taking into account metadata, full text of content, content descriptors, group classification, taxonomical classification, content descriptors including distribution details, use data, etc. In the query, the results matching the different contributions, different type of content, current user language are differently weighted.

- **Recommendations**: this module collects a large number of information to produce suggestions/recommendations. Suggestions are used to simplify the content search, and to push the users to access to new content, and to read news. The presentation of proposed lists of recommendations or suggestions or random selections from them is performed in real-time according to the user‘s accesses to the portal.
- **Mobile Support**: a module to provide web interface to mobile devices such as: i-Phone/iPad/iPod, Windows Mobile devices and other smart phones. It provides support to access at the collaborative activities provided by the Collaborative Support.

As seen in the modules description each of them includes a set of tools, services and functionalities, for this reason it is therefore not easy to make a detailed comparison with the four SECI modes of knowledge conversion: as regards the BPN modules, there are not always net boundaries between their features. However it is possible to make a comparison taking into account the main features for each BPN module (see Figure 3.5).

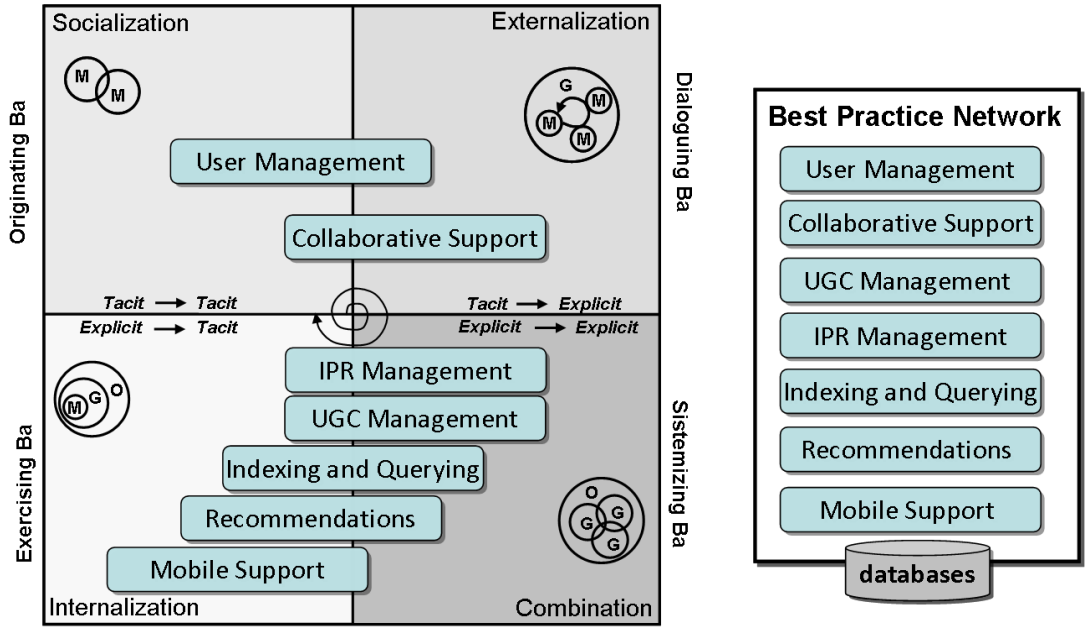

Figure 3.5: Match SECI – BPN Modules.

3.2.3. BPN Semantics computing

Semantics computing technologies may be profitably used to enforce more intelligence and efficiency in some of the services mentioned in the previous paragraphs. Thus, integrating technologies of content distribution with semantic processing and making decision support capabilities. In many applications, the semantic computing (as a support at the decisions, via the exploitation and processing of descriptors and semantic information) is confined on the server side to provide recommendations and reasoning about semantics, content and use data, users profiles, etc. Among these applications, collaborative solutions, social networks and recommendation systems are the most diffuse cases in which server side semantic computing is applied in several different extents, to provide a set of features and services to the users, [BeBFNP].
Most of the widespread Social Networks, are mainly focused on simple content formats (e.g., [YouTube], [Flickr], [LastFm]), [KrNa], while others are focused on: establishing social relationships among users (e.g. [Facebook], [MySpace], [Orkut], [Friendster]) [NaRaCh], searching users sharing same interests or people already acquainted with in the real world, and modelling groups [HRTS]. The former Social Networks organize and classify content on the basis of simple direct keywords, so that users access, retrieve and share them. Despite massive success of Social Networks in terms of penetration, most of them have limited semantic computing capabilities and provide simple recommendations about possible friends and for marginally similar content items.
Most of them can manage only simple audiovisual content (e.g., images or video, which are the simplest content items to be generated by users). On the other hand, the technologies of semantic computing may make reasoning about many other aspects such as: content descriptors, user profiles, device capabilities, use data, contextual data, etc.

The Best Practice Network manages and collects information to perform semantic computing and reasoning on the back-office and on the client side Content Organizer for indexing them and perform intelligent queries based on similarity distances, fuzzy search, and for computing suggestions/recommendations on the basis of static and dynamically collected information. Among the main semantic information managed:

- **user profile descriptions** collected statically via user registration and dynamically on the basis of user actions, migrated also on the mobile: selected content, performed queries, preferred content, suggested content, etc.;
- **relationships among users/colleagues/friends** (connection among users similarly to friendships, group joining) that impact on the user profile and are created via registration, by inviting other BPN users to joining them, performing registration to groups, etc.;
- **user groups descriptors** and their related discussion forums or blogs, multilingual web pages, with the possibilities of providing comments and votes (with taxonomical descriptors and full text in multiple languages for web pages and groups);

- **content descriptors** for simple and complex multilingual content (intelligent mini applications), web pages, forums, blogs, comments, technical descriptors, etc.; device capabilities for formal description of any acceptable content format and parameters, CPU capabilities, memory space, SSD space;
- **votes, comments, annotations on contents, forums, blogs, web pages**, etc., which are dynamic information related to users, content, forum, groups;
- **lists of elements marked as preferred by users,** which are dynamic information related to users;
- **detailed descriptors about downloads and play/executions** of simple and/or complex content on PC and mobiles, to keep trace of user actions as references to played content, which are dynamic information related to users preferences such as: date and time, GPS position, object ID, etc.;
- **descriptors about content uploaded and published by users** (only for registered users, and eventually supervised by the group administrator). Each Content element has its own static multilingual metadata, descriptors and taxonomy; while the related action of upload is a dynamic information associated with the user who performed it. In addition, content elements can be associated with groups.

3.2.4. Activities and Tools

In the BPN back-office a set of activities are performed to produce and/or to exploit the above-mentioned semantic information. Some of these activities produce data exploited by the front-end, to allow other activities to be performed in real time (i.e., on demand) on the basis of the users' requests. For example: queries on the multilingual indexes and semantic classifications and comments, the search of users on the basis of their profile, the identification and re-production of suitable content on the basis of device profile (web, mobile, PDA, etc.). On the other hand, there are many other back office activities which are computational intensive and that may take hours and days to be completed, and thus they have to be performed off-line, periodically, incrementally and sporadically in the presence of some specific conditions/events. The computational complexity may be due to one or more of the following causes, for example: content size (e.g., adaptation of real video files), number of users to be processed for estimating similarities (for example for clustering and recommendations), number of content items to be ingested and/or transcoded, [BBCFNP].
Among the several complex back-office activities the most relevant depict the content and information life cycle, and are those reported in the following list:

- **Content ingestion.** The users may upload new content via the portal and perform a semantic tagging on content elements (that is a part of the User Generated Content, UGC, workflow), this information has to be pre-processed to make it accessible for all the client platforms. Moreover, the content has to be also automatically ingested from archives (for example

those provided via OAI-PMH protocol), from databases, from file collections or from other means;

- **Content enrichment**. The several kinds of content in the portal provide a limited number of information for their indexing, so that they need to be enriched in order to provide enough information. Content items in the portal can be audio, video, applications, web pages, comments, forum topics, etc., and provide descriptors, multilingual information, taxonomical classification, association to groups, user ranking, comments, preferences, etc. In most cases, the content arrives without technical description, with limited multilingual data, etc. Some of this information has to be produced in multiple languages. The content enrichment may be strongly computationally intensive and may include:
    - **extraction of technical descriptors** for any digital resources which are needed for content adaptation and repurposing. The information is different from simple and complex content resources. Basic information may be: duration, resolution, size, dimension, video rate, sample rate and size, file format, MIME type, number of included files, file extension, etc. Different libraries or tools may be used to extract this information: FFMPEG for video and audio, SUPER, ImageMagik for images, etc. No one of them can be directly used for automated massive processing on multiple resource types and multiple computers.
    - **extraction of semantic descriptors.** The content can be also processed to extract information from the content data itself, for example: by processing images to understand the represented content (what is included into the image); documents to take a summary and extracting keywords; video files to segment major scenes and understand them; audio to extract tonality, rhythm, etc. These activities are strongly computational intensive. In this case, content semantic descriptors may be formalized in some XML, RDF and/or MPEG-7, and may be used to estimate similarity among content items.
    - **multilingual translation** of textual content and/or descriptors. The content typically reaches the archive and portal in one or a few more languages. An automated draft translation can be useful to make the users work the faster and simpler, and limited to validation. To this end, several translation services can be used.
- **Multilingual content indexing**. The several content types managed in the portal have to be indexed in order to provide fast search service to the multilingual users, taking into account their semantics. The multilingual version of textual information (metadata, descriptors, full text) has to be taken into account to perform a full multilingual indexing; in which, the several versions of information provided in different languages are referred to the same resource and not as in many portals in which each single language creates a separate index limited to the accessible translated items. Thus, main activities are those to cope with:
    - **indexing**. The indexing of each single language has to adopt a specific analyzer to take into account the grammar and lexical aspects of each single language. The insertion of the single

information into the database and into the Lucene indexing can be performed by the back office at each content insertion/upload, according to language analyzers. The index is performed at multiple levels: (1) performing multilingual translation and validation for accessible languages; (2) performing a multilingual text processing and analysis for fuzzy indexing on multiple languages [PiTa]; (3) collecting relationships and multilingual metadata and information in the database according to groups, taxonomy/ontology domain; (4) producing an integrated multilingual index of the information collected and optimizing the indexing for frontal multi-language search. The queries can be performed on the portal by using the above-mentioned modalities as simple text queries or as advanced queries specifying the modalities.

    - **taxonomy and ontology domain**. In a BPN, the problems of classification and meaning are relevant. A single content item may be assigned on multiple taxonomy nodes in multiple languages. Moreover, each single content may be described according to all its ancestor taxonomical descriptors. This creates a significant complexity when a node is matched in the query, all content items associated with the node and those of its children have to be candidates as well. Multiple paths can be obtained and thus duplicates in the query results, contributions coming from several paths may be collated by using different weights. On mobile devices an analysis and pre-processing have to be performed to make a smart indexing and thus the production of query results in real time. The produced information is collected on the BPN database to be processed and integrated into the single and multilingual indexes. The BPN developed can be applied in various contexts, as detailed in the next chapters (e.g. medical, performing arts, business, research, etc.) and the taxonomical classification has to be created and adapted basing on these context.

- **Content post-production and publication.** Once the content is ingested, it has to be processed to make it ready for the publication on the front-end distribution portal, filling all correct fields into the corresponding databases. Thus, a certain number of activities have to be performed in the back office such as:

    - **content adaptation.** The content, once uploaded/provided, needs to be adapted/pre-processed according to the portal needs and to the final user devices (PC, i-Phone, Windows Mobile, etc.). To this end, several different formats have to be produced according to the content type: audio, video, document, image, cross-media, etc. This activity of content adaptation and formatting is too computationally intensive to be performed on demand. Thus, an off-line pre-processing reduces the complexity at the expenses of the disk space.

    - **content packaging.** The packaging is performed to prepare the content for distribution by aggregating several kinds of information (e.g. digital essences and descriptors). This activity also simplifies the propagation of semantics information from portal server to

mobiles (e.g., suitable formats can be MXF, MPEG-21, AXMEDIS, MPEG-4, SCORM). The packing may include the application of some encoding algorithm, for example: the encryption to protect the content package or the single essences.

- **content publication.** It consists in the exporting of content and semantics to publication databases, posting the data on other social networks or portals, publishing them on P2P networks, for example on a P2P network based on BitTorrent, [Axmedis]. Aside to this activity of publication, the portal responsible needs to have the clearance on a number of rights for the publication of the information and data which is going to be publicly released. So that, this also means that the data/descriptors used to formalize the IPR Management and distribution portal behaviour have to be passed to the front-end portal.

- **Estimating suggestions** implies to perform a number of semantic computing activities that range from:
  - (i) estimating similarities among non-homogenous entities (e.g., symbolically and numerically described), among users' descriptors, user and content descriptors, content and content, taxonomy terms, annotations, etc.;
  - (ii) producing suggestions and recommendations by using techniques based on queries and/or clustering. The formal similarity among non-homogenous can be used for direct estimation of distances and as a basis of clustering techniques as described in the following.

The above-described activities of semantic computing are performed in the back office, which has been realized by using a scalable architecture based on AXCP grid [BeBN06]. The AXCP grid can manage parallel executions of processes (called rules) allocated on one or more computers (grid nodes). The rules are managed by a grid Scheduler (AXCP Grid Scheduler) and are formalized in extended JavaScript. The Scheduler performs the rule firing, resource and node discovering, error report and management, fail over, etc. It may put rules in execution (with parameters) periodically or when other applications request. It provides reporting information (e.g. notifications, exceptions, logs, etc., see Figure 3.6). The control and activation of rules can be also performed via a Web Service call by any application, such as workflow tool (such as Open Flow and BizTalk), portals in PHP, JSP, etc. This allowed the integration of the AXCP grid for content management and semantic computing with the Best Practice Network. The single grid node could invoke the execution of other rules by sending a request to the scheduler, so as to divide a complex rule into sub-rules running in parallel and use the computational resources accessible on the grid. An AXCP rule may perform activities of content and metadata ingestion, query and retrieval, storage, semantic computing, content formatting and adaptation, extraction of descriptors, transcoding, synchronisation, estimation of fingerprint, watermarking, indexing, summarization, metadata manipulation and mapping, packaging, protection and licensing, publication and distribution. AXCP nodes have plug-ins or may invoke tools to expand capabilities with customized/external algorithms and tools.

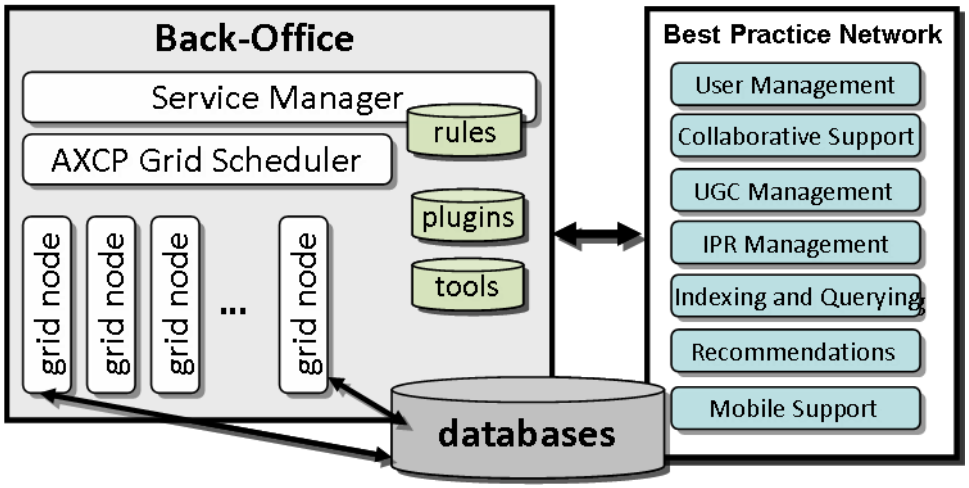

Figure 3.6: AXCP Grid for scalable semantic computing.

3.2.5. Recommendations and Suggestions via Semantic Computing

In this paragraph, the details regarding the production of suggestions among *elements* such as: users, content and user groups, are reported. Among the possible combinations only some of them are viable as described in **Errore. L'origine riferimento non è stata trovata.**. The suggestions have to be computed on the basis of relationships U→U, G→C, C→U, etc. where U means User, G: Group and C: Content, thus C→U means proposing Content suggestions to Users. Similarly, other suggestions can be managed for other elements such as: mailing lists, play lists, etc.
In this BPN the difference between suggestions and recommendations is the following:

- **suggestions** those computationally provided by the system;
- **recommendations** those produced by the single users, when they recommend a content/user to another friend, colleague.

		Recipient of the suggestions		
		User	**Content (played by a user)**	**Group (leader or members)**
Suggested elements	**Users**	Proposing to a user possible colleagues/friends	--- no sense ---	Proposing to a group leader possible interested colleagues to be invited in the group
	Contents	Proposing to a user possible interesting contents	Proposing at a play of a content similar content items	Proposing at a group member possible interesting contents (not much different with respect to C-C combination)

	Groups	Proposing to a user possible interesting groups	Proposing at a play of a content possible interesting groups in which similar contents are discussed	--- no sense ---

Table 3.1: Possible suggestions among elements.

The earliest solutions for guessing users' intentions have been based on keyword-based queries (i.e., sponsored search, or paid listing), which places ads and/or suggestions in the search results; and content match, also called content-targeted advertising, or contextual advertising, which places ads on the basis of the Web page content and content similarity [CiMuPl], [AnBrGaJo]. The latter is a harder problem with respect to the former.
Contextual suggestions are widespread and many systems are capable to extract keywords from web pages for producing suggestions [YiGoCa], [ZhSuPlNa], sometimes using semantic approaches [BrFoJoRi]. In order to predict which terms are more relevant to a set of seed terms describing a product or service, models based on clustering, collaborative filtering, logistic regression, etc., are used, as in [BaMuSe].
Suggestions can be computed by using several different techniques by estimating similarities, making queries on keywords, adopting clustering or grid clustering, etc. In most cases, the elementary operation for estimating suggestions is the similarity distance between elements. When similarity has to be estimated on the basis of simple profiles/descriptors may have a limited complexity. On the other hand, the estimation of all distances or similarities among elements can be computationally expensive in the presence of complex descriptors and/or millions of items, since it can be quadratic, an $O(N_2)$ (where N is the number of elements) or much more taking into account the complexity of the descriptors that may be based on several fields.
Thus, the problem of providing suggestions is reduced to the estimation of similarities among elements, which in the BPN are:

- **Users** of the portal with information collected from users:
    - **static** aspects generically provided during registration. The static part is frequently not so much detailed in generic Social Networks, since users prefer to avoid filling in "useless" forms and/or to provide false data. In small thematic and business oriented Social Networks the information is much more reliable. Among them:
        - general information: name, surname, nickname, gender, age, location, foreign language, mother tongue, nationality, etc.;
        - instant messaging contacts: skype, messenger, ICQ, etc.;
        - education and job, interests: skills, content Type and Format, or taxonomy;
        - role covered in the platform (e.g. teacher, group manager, platform administrator, etc.) and the access permissions that correspond to the role (related to: contents, discussions,

groups, functionalities, etc.). It has to be noticed that the type of the user roles available depends on the context into which the BPN is applied;

- **dynamic** information collected on the basis of the activities that the users perform on the portal elements, such as those on content or on other users:
  - votes and comments/annotations on: contents, forums, web pages, etc.;
  - downloads and play/view/executions of content, web pages, etc.;
  - uploads and publishing of user provided content;
  - mark content as preferred/favourite;
  - recommendations of content/groups or users to other users;
  - chat with other users, publish on groups;
  - queries performed on the portal, etc.;
  - create a topic in a forum or contribute to a discussion;
  - relationships/connections with other users or groups;

- **Groups** of users they may have specific descriptors and those inherited by the users:
  - **static** description of the groups such as:
    - objectives, topics, web pages, keywords, taxonomy, etc.;
  - **dynamic** aspects related to:
    - users belonging to the group; users may: join and leave the group, be more or less active over time;
    - content associated with the group: files, comments, etc., with their taxonomical descriptor, metadata and descriptors.

Thus, most of the descriptors may change over time (for example how they may depend on the user's actions: votes, play, download, comments, user joining a group, etc.) and therefore, distances and similarities should be updated at each change, even in real time. On the other hand, when it comes to cross media content, it can be very complex since you could shift from a simple classification into a multidimensional semantic classification. As to cross media content, what is meant is content which may have multiple format (audio, video, image, document and cross media). Thus, the cross media format files may include a collection of other contents glued together with a presentation layer based on SMIL, HTML, FLASH, ePub, and/or MPEG-4 BIFS and all the related descriptors.
Moreover, when a new element is added (a new user, or group, or content, etc.) the estimation of a significant number of distances could be needed to both: providing suggestions and to take into account of it for the suggestions to be provided to other elements.
The estimations of the new distances can be limited to the new elements added, G, with respect to those which are present, M, so that the computational complexity can be limited to an $O(GM)$, when $M >> G$.

On such grounds, it is evident how the costs for the semantic computing of suggestions can be very high. Solutions to reduce complexities are based on clustering techniques, grid and progressive clustering, and incremental estimations of similarity distances among elements. When the complexity of element descriptors becomes high, as it occurs with cross media content or in the presence of complex descriptors, the problem of complexity management gets more evident. The basic problem is to model the similarity estimation among heterogeneous elements forming the description of an element, so as to guarantee any possible estimation even among elements of different kind and in the presence of uncertain and/or incomplete data. One solution to reduce the problem is to identify the minimal number of features (descriptor aspect) which are significant. This can be performed by using principal component analysis, PCA, or in any case by means of statistic analysis.

3.2.6. Clustering Techniques

A solution to reduce the computational complexity of suggestions is the adoption of clustering techniques, which allow to group elements in families, for example by using *k*-means, *k*-medoids, hierarchical clustering, etc. This allows to reduce the complexity from *N* elements to *K* clusters, where *K* is typically much smaller than *N*. The suggestions are thus provided by estimating the similarity between the recipient of the suggestion and the closest cluster descriptor. Thus, the effective suggestions may be directly provided by randomly picking elements from the most similar cluster, or by estimating specific distances/similarities on a smaller set.
Once the clusters are defined, they depend on similarity distances, which have static and dynamic parameters. It means that at each change of dynamic parameter or when a new element is added, a new clustering estimation would be needed. This can be avoided by applying the estimations only periodically and limiting the estimations to the new distances and to those that have been invalidated by changes. The new elements can be initially joined in a cluster according to its proximity with the descriptor representing the cluster (for example the cluster center).
Different semantic modelling and clustering techniques can be adopted, here after the main studied for the BPN developed are synthetically described:

- based on the **k-means algorithm**. The k-means clustering unsupervised method is in most cases based on the possibility of estimating direct distance among elements. It provides quite good performance in terms of scalability; discovery of clusters with arbitrary shape; minimal requirements for domain knowledge to determine input parameters; ability to deal with noise and outliers; insensitivity to order of input records; and support for high dimensionality, it has typically a complexity of an O(NKI), where N is the number of elements, K the number of clusters and I the number of iterations. Many other clustering algorithms exist, while the k-means has demonstrated the best performances when N is largely bigger than K and I [EvLaLe];
- based on the **k-medoids algorithm** the k-medoids and/or hierarchical methods (see [EvLaLe], [XuWu]) and finally a hierarchical evolution of the k-medoids. K-medoids is a partitioning technique which is based on

the fact that one of the elements (called medoid) would be centrally located for each cluster.

- **hierarchical clustering**. This clustering method creates clusters on the basis of the distance among the single elements, [XuWu]. The process starts by aggregating the closest elements to create smaller clusters of two elements and then aggregating these small clusters with other by following a sort a merging algorithm. The aggregation is based on the distance metric such as those discussed in the sequel of the chapter. Hierarchical algorithms may differ for the mathematical model used for the merging of sub-clusters: complete linkage, single linkage and averaged.

3.2.7. Similarity Distances

According to Table 3.1 and the descriptors adopted on the Best practice Network, **Errore. L'origine riferimento non è stata trovata.** reports the similarity distances used in the clustering model and to estimate suggestions. It has to be noticed that, despite to the fact that Users, Content and Groups may all present both static and dynamic aspects, the dynamic aspects have been only taken into account for similarities involving Users. This has limited the complexity, while it should be noted that the metrics reported in this chapter can be used to extend the estimations to take into account dynamic aspects of Contents and Groups which are related to the activities of Users on them. For example, the similarity distance D(C,U) depends on both static (s) and dynamic (d) aspects: D(C(s); U(s,d)).
It may be used to generate suggestion C→U taking into account the profiles of preferred contents of the Users, the dynamic information collected on the basis of the content play, favourite, highly ranked, etc., with respect to the content descriptors. In short, the similarity distance is practically reduced to estimate the distance between a content and the user represented by the descriptor of content he/she: marked as preferred, positively ranked, uploaded, played, recommended, etc. Please note that preferred content description is a dynamic aspect of the user that presents information that can be compared with the static descriptions of the Content, for example in terms of taxonomical multilingual classification.

		Recipient of the suggestions		
		User	**Content (played by a user)**	**Group (leader or members)**
Suggested elements	**Users**	D(U(s,d); U(s,d))	--- no sense ---	D(U(s,d); G(s,d))
	Contents	D(C(s); U(s,d))	D(C(s); C(s))	D(C(s); G(s,d))
	Groups	D(G(s,d); U(s,d))	D(G(s,d); C(s))	--- no sense ---

Table 3.2: Elementary Similarity distances.

In general, the above-mentioned similarities distances can be scalar or vector data. For example, a scalar model may be obtained for the *D(U,U)* taking into account both dynamic and static information:

$$\mathrm{D}(\mathrm{U1};\mathrm{U2}) = k_s(\sum_{i=1}^{\mathrm{T_s}} x_i \mathrm{Sd}_i(\mathrm{U1},\mathrm{U2})) + k_d(\sum_{i=1}^{\mathrm{T_d}} y_i \mathrm{Dd}_i(\mathrm{U1},\mathrm{U2})) \quad (1)$$

Where:

- Sd_i and Dd_i are the distance metrics for a static and dynamic feature, respectively (static distances can be computed once);
- k_S and k_d are weighting the static aspects with respect to the dynamic aspects and to adjust the scale factor among them according to the number of metrics and their ranges (they have to be fixed on the basis of the portal intention, on the other hand the dynamic aspects are much more reliable than static as previously commented);
- T_s, T_d are the number of static and dynamic features to estimate the similarity distance, respectively.
- x_i and y_i are the weights to give different relevance to the corresponding feature metrics.

The vector model leads to keep separate the single metrics:

$$D(U1;U2) = \begin{Bmatrix} k_s(x_1 \mathrm{Sd}_1(\mathrm{U1},\mathrm{U2}), x_2 \mathrm{Sd}_2(\mathrm{U1},\mathrm{U2}), \ldots, x_n \mathrm{Sd}_{Ts}(\mathrm{U1},\mathrm{U2})) \\ k_d(y_1 \mathrm{Dd}_1(\mathrm{U1},\mathrm{U2}), y_2 \mathrm{Dd}_2(\mathrm{U1},\mathrm{U2}), \ldots, y_n \mathrm{Dd}_{Td}(\mathrm{U1},\mathrm{U2})) \end{Bmatrix} (2)$$

Both approaches can be used in clustering techniques presented above. For example, in the case of *D(U,U)* distances, a reduction of dimension could increase the computational cost since users presents a large set of metadata. In both cases, the single metric can be separately weighted more with respect to the others or they can be weighted per groups, for example, to give more relevance to dynamic aspects rather than to static. The weights of the above formulas (1) and (2) can be estimated by using a multi-linear regression by considering the goals of the portal organiser or of the community of users [RoLe], [FioNe]. These techniques are used in the area of empirical assessment and evaluation.
In both cases, the problem is the definition of the single metrics. Thus, in the next sections, some of the Static (Sd_i) and Dynamic (Dd_i) metrics are described. They constitute the elementary blocks for the estimation for the similarity distances reported in **Table 3.2**.

3.2.8. Content Aggregation

The BPN developed includes functionalities related to the e-learning and to the collaborative work activities, in particular provides an integration with the LCMS Moodle, that is one of the most diffused Content Learning Platforms Systems, [Moodle]. One of the fundamental aspects for the conduct of a lesson or more generally to collaborative work, is the opportunity of creating aggregations of contents. These aggregations can occur with various modalities and on the basis of both the characteristics of the content and the needs that users themselves have in a given context.

In the BPN, several different kinds of contents and data are modelled ranging from content to users and their relationships, moreover the different kinds of contents and also the users are associated with a thematic taxonomy which describes the contents in terms of its features (e.g. genre, historical period, subjects, management aspects, dissemination aspects, etc.). Also the user profile includes such a classification to allow users to express their preferences about content theme, [BeNePaSeDMS]. Taking into account all these factors, three different modalities to realise the content aggregations have been studied and implemented:

- **Playlists**: a play list can be a set of audiovisual contents. Play lists are a way of playing a series of multimedia contents in a specific order, defined during the authoring phase. The play list in the BPN developed, is not the classical play list of many social network, since in ECLAP the user may chose to put in the play list not an entire video or audio, but a segment of video, audio or an image for a number of seconds. This can be done through a proper player that will be loaded during the editing phase in which user can select media, add them to play lists box, choose an order and choose a segment for each media. Once the edit phase is complete, user can save the play list assigning full metadata to relate, in order to save it as an object that will be stored in database, will be available for the playback session, and will be shared with other users. Thanks to the metadata, this object will be indexed properly, to keep easier the search on the portal.
- **Collections**. A collection is a set of contents: audio, video, images, documents, pdf, play lists, zip files, etc., that are grouped together according to the personal purpose of the user. They may be thematic collection as well as used as the first step to collect content for preparing a lesson for e-learning environment. The teachers use the Collections as sources of contents that they can be grouped together following a topic or by choosing a connection amongst them. The Collections are directly created on the Portal and may be published or kept private by the user. The collections are automatically exported, on the LMS side of ECLAP to be used by the teachers to create a lesson.
- **Semantic Annotations**. MyStoryPalyer is a tool that allows users to be the central part in fruition of multimedia objects annotating them, and offers new solutions for educational and infotainment purposes. The innovative part of MyStoryPlayer lies in the fact that no difference between media and the user's annotation exists, because both categories are referred to multimedia objects and they are temporally connected. Annotations are related to one or two media, and each media can be associated to many annotations. Therefore, according to the model, two audiovisual objects can be synchronized each other through an annotation containing temporal elements that define this kind of relation. An annotation involves one or two media and is composed by a start time, an end time, an identifier of the media annotated, an identifier of the media to relate with the first one, and a textual description of the annotated segment. This information are codified by RDF triples and saved to a Sesame database, external to the BPN one, interfaced with the Flash MyStoryPlayer that interprets the information and play the contents in synchronized way according to their temporal lines. As for the other

features like play list and collection, no alteration is done on the original file, when two media are synchronized, because the system provide to generate new code RDF that identify this relation, and save on internal database the information about the annotation in order to allow the user to access to his annotations from his profile.

3.2.9. Metrics for Best Practice Networks Analysis

BPNs organize and classify content, so that users can easily access, retrieve and share. In the BPN developed, an engine to monitor data has been realized in order to reason upon content descriptors, user profiles, ads profiles, user generated content descriptors, device capabilities descriptors, etc. To keep trace and analyze user activities is becoming more and more fundamental for tuning services and predicting needs, such as market trends, reactions of users, the product's acceptance, educational needs and problems, [BCFNPPa]. In order to provide the Portal administrator with an effective tool able to identify the general trends of the BPN including downloads and user behaviour, a set of metrics has been defined. Therefore, the identified set of metrics has been implemented. The selected metrics are the most relevant:

- **Usage of Devices and Platform**: as mentioned, the BPN offers a wide range of contents may be accessible by different types of devices (PDA, i-Phone, etc.) and platforms (Linux, Mac Os X, etc.), in order to improve the BPN developing it is useful to monitor what are the most platforms or devices used;
- **Proposed and Preferred Content Formats**: given the great diversity of content the platform offers, it was necessary to identify which kinds of format are more appreciated by the BPN users with respect to those provided within BPN. It is also interesting to compare the content offered with those most downloaded by the users, in terms of formats;
- **Content Access Modality of Users.** The content accessed via PC and mobiles may be downloaded or played via the HTTP connection in progressive download. This last modality is possible on PC and i-Phone, while presently it is not viable on Windows Mobile PDA devices endowed with AxObjectFinder tool. On i-Phone, it is possible to have online execution of content or download and successive offline play;
- **Object Stability**: In a BPN the distribution of the content accesses is exponential but stable. On such grounds, the Object stability is an interesting metric that allows to understand which is the level of interest for a given content item within the BPN community. This metric measures how many times a content object has been downloaded in a period of time;
- **User Fidelity**: The aim of User Fidelity measure is to assess in a simple shot which are the most active users in the BPN. The measure takes into account the number of active days in the period for each user. This simple metric is strongly correlated to the number of content: preferred and uploaded.

The metrics and statistics that have just been described in summary provide information related to the performance of the platform, they can contain sensitive data, are fundamental to develop the platform and improve its features and functionalities taking into account the users behaviours: for all these reasons they are accessible only by the site administrators. In order to better comprehend the relevance of these statistics, real data are used in the "Mobile Medicine" scenario (see chapter 4).
However some statistics can be useful also to stimulate the user to produce more actions on the portal, to better use the tools provided to them, to connect and communicate with other user of the platform, etc., so other statistics have been realised and can be visualised by all the registered users or by the group managers. The most relevant are listed below:

- **Personal statistics**, accessible only by the single user and by the Administrator, on:
    - **content upload** (e.g. number of contents uploaded by the user, number of contents that have been placed as favourite from other users, profile of the content uploaded the user in terms of taxonomy items, number of contents that have been placed as favourite from other users, etc);
    - **content access** (e.g. number of the downloads of the contents that the user has uploaded, temporal distribution of the play/download of the content that has been uploaded by the user, number of the total play/download made by the user, profile of content that the user has played/downloaded in terms of taxonomy items, temporal distribution of the preferred formats of the content played/downloaded by the user, etc.);
    - **networking** (number of the comments made by the user on forums, number of the forum topics created by the users, etc.);
    - **content enrichment** (number of times a specific tool has been used/opened to: modify metadata describing the content, make annotations, comment or votes on the content; number of metadata changes, number of annotations, number of collections or play lists, etc.);
    - **social promotion** (user eccentricity value, user betweenness value, number of connected colleagues/friends, Number of people invited to visit the platform, etc.).
- **Group statistics**, are accessible only by the group managers. Each group manager can see a set of statistic for each of the group managed (number of contents uploaded into the group, preferred format uploaded in the group, total number of accesses in terms of play/download of the group contents, temporal distribution of accesses in terms of play/download of the contents that have been uploaded in the group from PC and PDA, etc.).

3.2.10. Social network Analysis

Once having registered data related to all Portal activities, content, users, etc. it is possible to proceed with the monitoring activity basing on Social Network Analysis (SNA). This phase is particularly relevant in order to make provisions for the best services that could be provided to users, to establish what users are more relevant (referring to graph theory), discover how to improve the aggregation activities, to create and provide to the users of the Portal new functionalities, that are based on their preferences, etc., [EC-DE3.1].

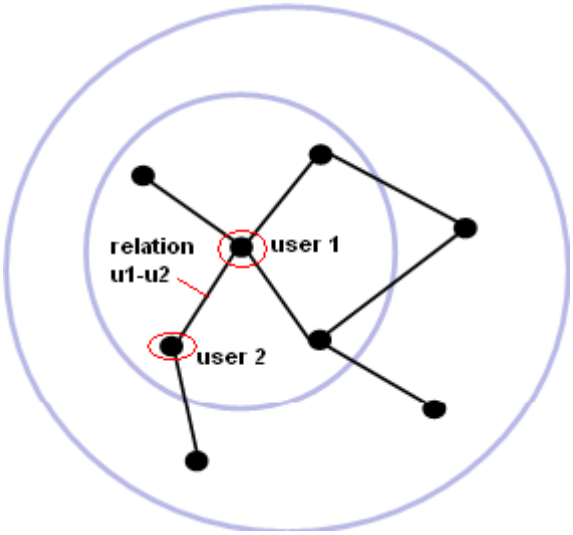

Figure 3.7: Network of relationships.

The SNA is mainly focused on evaluating the status of the network of relationships that may give an idea of the evolution of the BPN and of the healthy aspects. In order to do this models and metrics to establish the following features are be elaborated:

- Connections among users basing on graph theory (the users the node on the graph and the relations among them are the edges, Figure 3.7):
  - Number of direct connections among users,
  - Distances (number of nodes-users) among users, groups, cliques, etc.
  - Which are the most important users
  - Which are the critical conditions
  - Which are the major drivers of growing
  - Which are the most interested aspects in a given period
  - Which are the most active people
  - etc.
- Determine user preferences basing on their activities:
  - downloaded content
  - query made to search content
  - Annotations on content
  - etc.

The main metrics used are:

- **User Eccentricity**: is defined as the max distance of a certain node with respect to all other nodes of the network.
- **User Betweennes**: gives info about the control degree of a node (user) about the information flowing among other nodes (users).
- **Number of connections**: number of friend for each user (direct connections)

4. Mobile Medicine

In many health care situations, powerful mobile tools may help to make decisions and provide support for continuous education and training. They can be useful in emergency conditions and for the supervised application of protocols and procedures. To this end, content models and formats with semantic and intelligence provide more flexibility to put in the hands of medical personnel (in off-line and on-line conditions) powerful tools with respect to those that are presently adopted and on market. In this chapter, the Mobile Medicine solution is described, it exploits a collection of semantic computing technologies together with intelligent content model and tools to provide innovative services for medical personnel. Most of the activities of semantic computing are performed on the back office on a cloud computing architecture for: clustering, recommendations, intelligent content production and adaptation. The mobile devices have been endowed of a content organizer to collect local data, provide local suggestions, while supporting taxonomical searches and local queries on PDA and i-Phone. The proposed solution is under usage at the main hospital in Florence. The smart content has been produced, by medical personnel, with the adoption of the new ADF-Design authoring tool, which produces content in MPEG-21 format. The mobile content distribution service is integrated with a collaborative networking portal, for discussion on procedures and content, thus suggestions are provided on both PC and Mobiles (PDA and i-Phone).

4.1. Medical Needs

Medical personnel need to access fresh information and knowledge in emergency conditions, and during activities of continuous medical education and training. This knowledge, support medical/paramedical personnel in the adoption of continuously evolving standards, intervention protocols, complex dosages for pharmaceutical prescriptions depending on the context and patient status, etc. In hospital, the needed information and knowledge regarding these aspects is continuously updated and have to be propagated in short or real-time; for example, via desktop terminals and in some cases reprinted on paper.
In a paperless hospital and in emergency/critical-conditions, mobile devices are mandatory tools for information access and thus to take important decisions; thus, the solution has to guarantee access to right and updated information in the needed time [PrGrHe], [MiRuSlWi].

In the Mobile Medicine scenario ([MobMed]), there is the need of performing semantics computing on mobile devices and in general on client side. Semantic computing capabilities on the mobile device can be useful to provide support for making faster decisions also in off-line conditions. For example, to take into account the context, to provide local and personal recommendations, to select and use dosage tools, to recover health care procedures, to perform classification and assessment models, to identify and follow intervention models, to access at suitable educational content, to access and exploit interactive technical manuals, etc.
Therefore, in order to integrate autonomous capabilities and semantic computing on mobile devices, intelligent content models and packaging solutions for delivering them are needed. They may have capabilities of enforcing/modelling multiple paths or experiences; exploiting complex semantics and descriptors, creating interactive intelligent content with semantic computing capability, and when possible taking into account of server side fresh information.

The comparison of the state of the art of Mobile Medicine system can be performed only at level of single technology since there are no other similar integrated solutions available. Therefore, it has been decided to design and develop a Mobile Medicine solution to cope with the above-mentioned problems exploiting the possibilities of semantic computing, in an integrated solution for the medical area providing support for:

- **delivering and exploiting on mobile devices intelligent content** capable to help making decisions, estimating dosages, performing assessments, collecting data, etc.;
- **providing on-line suggestions** via semantic computing on servers of the web portal, automated/assisted content production, collaborative content finalization and semantic enrichment;
- **providing off-line suggestions** via local semantic computing on the mobile devices.

Mobile Medicine is presently under trial at the largest Florence Hospital joined with University of Florence health care area (i.e., Careggi Hospital). It is functional for desktop computers (i.e., Microsoft Windows, Linux, MAC OS) and for mobile devices such as i-Phone/iPod/iPad, and PDA as Windows Mobile based phones. The most innovative aspects are related to the automated production, semantic enrichment, distribution and exploitation of complex content and thus to the exploitation of semantic computing capabilities on PC and mobile devices, thanks to the distribution of packaged content. Among the semantic computing aspects: indexing and querying, automated adaptation, user and content recommendations.

4.2. Mobile Medicine Requirements

According to the above described scenario, the most important requirements have been identified. The most relevant capabilities are related to provide support at the medical personnel, (see Figure 4.1, [BBCFNP]):

- **during emergency and critical conditions (on-line and off-line)**, to get right content that may be useful to make decisions such as: estimating dosages, estimating objective assessment models, guiding on choosing medicine and solutions, reminding precise procedures/protocols, reminding the usages of accessible instruments (for example in an ambulance), etc. This requirement may be satisfied by using a set of technologies: server side recommendations, local recommendations, intelligent content model, suitable servers exploiting semantic computing on content, off-line availability of these features, etc.
- **for Continuous Medical Education/formation via mobile devices (CME)**. CME means lifelong education of medical personnel, this implies to define courses, assign them to personnel, and monitoring progresses, etc., using mobile devices. The adoption of mobile devices for medical education is still not much diffuse. Moreover, in this context we intend education focused on the exploitation of the intelligent content and tools.
- **for producing and refining intelligent content with simple assistive tools**. This feature means to provide a service to (a) collect a range of possible media types and automatically make them suitable for a range of devices, (b) provide production tools for producing intelligent content for decision support (as described above) including the production of semantic descriptors and packaged content for distribution. For point (b), production tools have to be intuitive and usable for non-ICT specialists such as the medical personnel. So that they can be capable to autonomously completing and refining content.

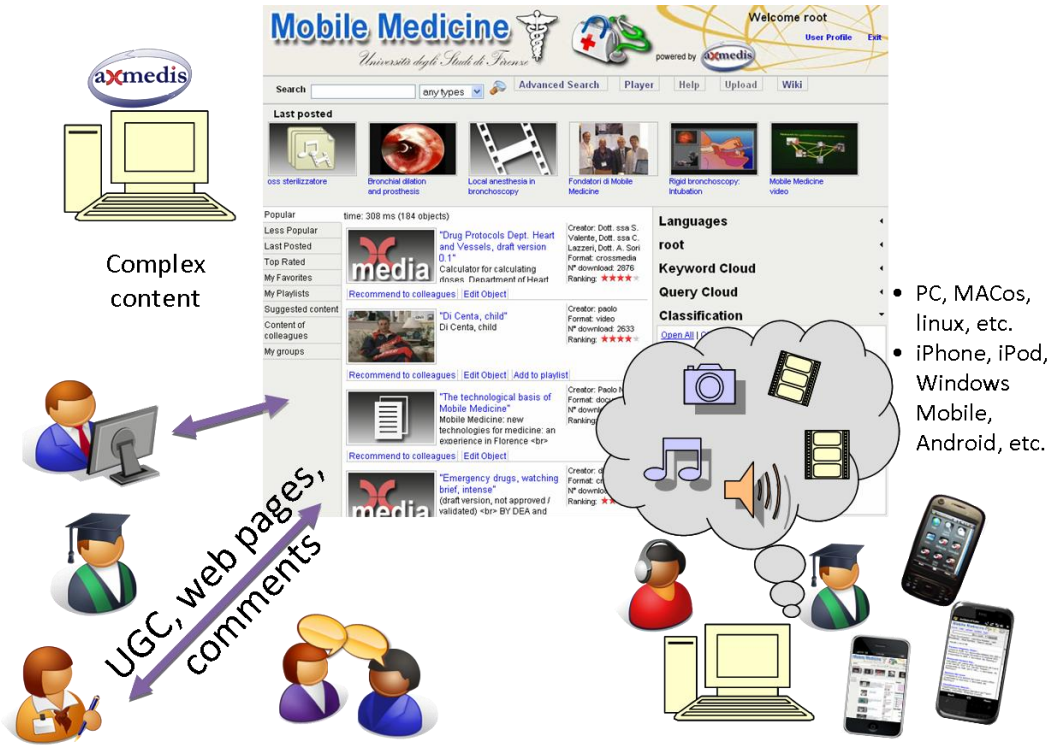

Figure 4.1: Mobile Medicine capabilities and supports.

4.3. Final Users Players and Tools

In this section it is fundamental to clarify what is meant for complex objects. They are not simple digital objects i.e. audio or video file, but real applications, for example: interactive calculators that work with specific formulas (proportioners, dose estimates), algorithms that help to follow the procedures, guided courses, sequences of slides, interactive games, animations, etc. These complex objects must be seen as a real mini-applications that can be performed using the AXMEDIS player on PC, PDA and/or even directly on the i-Phone or on the browser you have. For the execution of complex objects (mini-applications) in certain cases it is necessary to make the installation of the AXMEDIS player (PC or PDA), or can be displayed in the Internet browser, [MobMedUM].
On Mobile Medicine portal, there are several examples of these mini-applications. Other examples of content/cross media applications can be obtained/downloaded from the portal XMF (Cross Media Finder, [XMF]). The format used for files is MPEG-21/AXMEDIS, [Axmedis].

In the Mobile Medicine scenario, the content types range from *single files*: audio, video, images, documents, slides and animations; to *cross media* containing supportive tool such as a sort of mini-applications (internally consisting of a set of files and descriptors). As depicted in Figure 4.2, examples of cross media content can be, [BBCFNP]:

- **interactive guidelines/procedures** to help users to remind the correct procedures and help them make decisions on the specific cases and pathologies. They are internally built similarly to flow charts describing what has to be done from general to specific cases (for example in emergency conditions), as described in the following. The medical personnel is guided to answer to a set of questions that lead the system to identify precisely the context and thus the actions to be performed on the patient and its general conditions. The so called triage model typically may underline the whole process with several other procedures connected to it to address specific cases;
- **calculators and tools** that are interactive applications in which the user may insert data/info collected from the patient/context (such as: age, weight, pressure, temperature, habits, symptoms, conditions, reactions, etc.) to obtain from the device/content some estimations/suggestions to be used to make decisions. For example, the estimation about the probability of pulmonary emboli, the estimation of a dosage on the basis of patient weight, the estimation of fat percentage, the suggestion of a prescription about what you could eat, the assessment of neurological conditions on the basis of standard quantitative models, etc.

The above-described content model has to provide a certain degree of autonomy to react at user stimulus and to provide support to make decisions. They may be composed or linked to create more complex content solutions and paths for both fixed and mobile devices and accessible off-line. To this end, the cross media content has be indexed in local database to be searched and thus activated in short time at disposal of the medical personnel using the mobile device, without having

the needs to wait for the web page nor for the content download. Thus, the cross media content may take advantage of the communication with the server if any, and may provide hints and suggestions to the user in any case, even when the mobile devices is off-line. For the above-mentioned complex content, an authoring tool is available to allow medical personnel to create their own procedures and tools, [MobMed].

Figure 4.2: Cross Media Intelligent Content of Mobile Medicine: video, audio, sliding, assessment, data collection, dosages, decision support.

At level for content format, the content/applications are modelled as AXMEDIS packages in binary/xml file format [BeBNR], [BeBN06], enforcing in addition Mobile Medicine semantic descriptors and capabilities. The format structure supports the navigation, the creation of nesting levels and the direct access to the resources via links and references. The interaction aspects in the model are delegated to the presentation layer. The binary version of the format includes a table for immediate access to digital essences even when they are encrypted and it is based on ISO Media file format [BeNR], [BeBNS]. This allows content to be downloaded, streamed and/or progressively downloaded and played in real time, even when the essences are protected.

In the Mobile Medicine format **multiple descriptors**, identification and classification info are hosted: Dublin Core, extended metadata and descriptors, taxonomical association and RDFS formalizations, plus AXInfo (a set of metadata used to manage the content life-cycle), workflow info; plus any other identification and/or descriptors in XML, and/or RDF.
Descriptors included in the content package may refer to internal and/or external digital essences. The format supports **multiple presentations** in the same package. For example, it may provide and alternating HTML/CSS/JS, XML, SMIL, ePub, NewsML, MPEG-4 and FLASH in dominating the main canvas of the player.

HTML and other XML formats may use style sheets and digital essences (text, video, audio, image, etc.) hosted into the content package. On AXMEDIS players, presentation layers (such as SMIL, HTML, and FLASH) and also internal events may put in execution some methods written in extended JavaScript for activating **behavioural actions** allowing to perform a range of functionalities of the intelligent content including: inspect and modify the content structure (e.g., add new resources, internal search, reflection), control the resources rendering, perform calls to web services, making decisions (on the basis of user profile, context, actions performed, descriptors, rule based, etc.), activating semantic computing, communicating with servers, recovering GPS position, recovering data from web, etc.

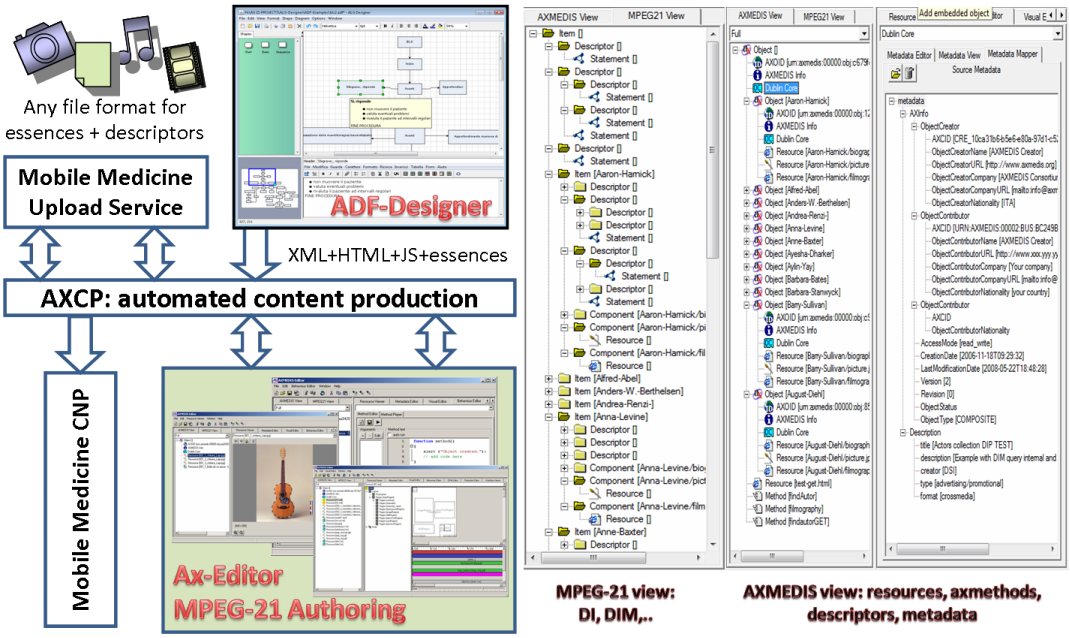

Figure 4.3: Content production flow and format structure in MPEG-21 and package views.

Thus, the content behaviour is specified by coding the business logic intelligence with JavaScript. Semantic computing and processing capabilities are activated via JavaScript by using the same formalism of the AXCP back office language grid mentioned before. Each single content may implement different communication modalities with the portal server to get and post information. The content may make decisions locally or may demand making decisions to fresh server information according to the communication conditions.
The production of Mobile Medicine content is performed by using a set of integrated tools (see Figure 4.3):

- **automated production** exploiting AXCP extended rules functionalities. When the UGC is produced by the prosumers and uploaded on the portal, a number of metadata are requested, while other descriptors are automatically added when digital essences are adapted and the package produced for publication on CNP (for PC and mobiles according to the IPR). This modality is used for single files: video, audio, images, documents, etc., in any format. During the upload, the prosumer provides some metadata, grouping and taxonomical classification which are used by the portal and on the smart phone for semantic computing. Technical and

other semantic descriptors are automatically added by the AXCP and included into the content package.

- **assisted production of complex intelligent content** (emergency procedures, questionnaires, making decision supports, etc.) by using a visual language as depicted in Figure 4, which shows the ADF-Designer tool. The ADF visual formalism has been specifically designed to formalize Medical procedures and it is simpler than the flow chart (screen pages connected by possible other successive pages). Details about the single screen are provided with a HTML graphic editor tool. The interaction capabilities are confined in HTML controls. The ADF tool automatically generates the integration and business logic code in a set of files, which is used to automatically produce the Mobile Medicine content package in MPEG-21 which can be uploaded on the portal. During the content ingestion a specific descriptor for propagating the semantics into the i-Phone Mobile Medicine application is produced.
- **manual authoring**. The AXMEDIS Editor is a graphic authoring tool for MPEG-21, including: structure editor, presentation tool editor (HTML and SMIL), metadata and descriptor editor, behaviour editor and simulator, workflow editor, IPR licensing tools, and protection tool on packager.

The Mobile Medicine content package is the vehicle to move semantic descriptors from portal to mobile devices and as the container in which the semantic computing capabilities find a context on which they can make decisions and interact with the general distribution portal. In fact, cross media content on mobile players executes content on which the content changes behaviour depending on user profile and actions, or on basis of local (GPS, accelerometers, device status, communication status, etc.) and server context (server files and info, server accessibility); it may communicate to the server use data or other information regarding the mobile status. The Mobile Medicine content may be used to actively and proactively: (i) collect patient data, (ii) remind at medical personnel or patients to perform certain actions or to follow checklists, (iii) collect annotations proactively, (iv) create multimedia scrapbooks, (v) collect user generated content, etc.;

4.4. Mobile Device Content Organizer for Local Semantic Computing

In the context of the Mobile Medicine, it is very important to provide a certain number of services in off-line conditions. The solution has to be user friendly, so that medical personnel do not have to take care about content/application life cycle. To this end, a Content Organizer (technically called ObjectFinder) for smart phone has been designed for the Mobile Medicine (it is freely distributed for PDA Windows Mobile, and for i-Phone) (see Figure 4.4, for the architecture of the PDA version, a similar structure is also present in the i-Phone version that you can download from Apple Store for free).

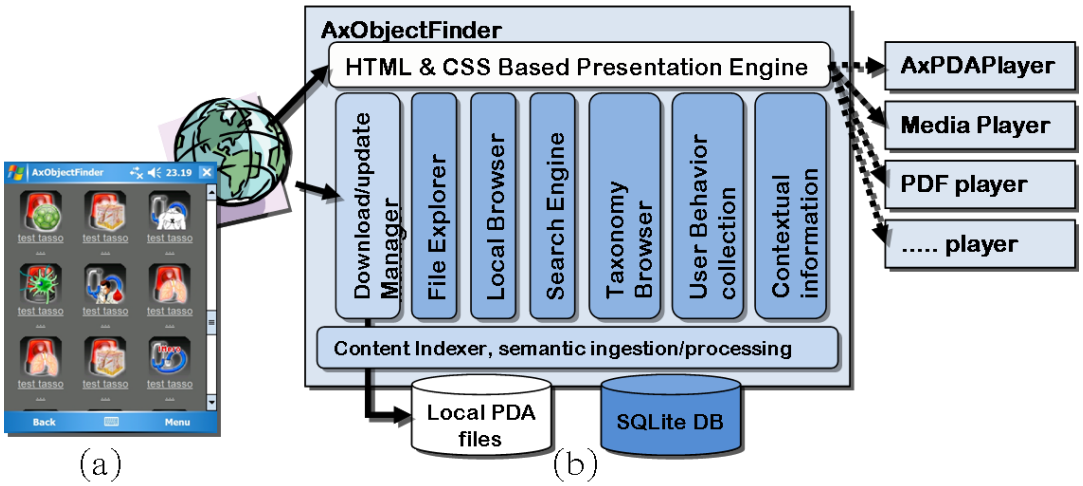

Figure 4.4: ObjectFinder for PDA: (a) user interface, (b) architecture.

The ObjectFinder presents a set of essential semantic computing capabilities; it processes the content items to index them according to the descriptors, taxonomical classification of medicine, and provides support for querying and organizing content according to the user data and requests. It has been designed to support the users to collect, index, organize, update, search and retrieve content on the basis of their semantic descriptors, user profile (static and dynamic, including use data), user preferences, etc. Thus, it enforces semantic computing and intelligence at level of mobile device and takes advantage of the management of intelligent content it can download and collect into the smart phone. The ObjectFinder provides a direct usable interface based on icons for main functionalities: portal access, local search, taxonomy based content browsing, suggestions, access files, etc.; and for content play with a single finger click. The ObjectFinder Presentation Engine is based on HTML with a protocol to access at objects stored into the mobiles device (e.g., those with icons and query results).

The Content Organizer ObjectFinder has the capability of:

- **collecting and indexing cross media content** coming from several channels: from the distribution portal, from the connected computers (e.g., USB, Bluetooth, IRDA), network connection (e.g., HTTP from a web page download). The indexing is performed exploiting MPEG-21 package metadata and semantic information into the local database of the mobile to exploit search/query/reasoning facilities for the final user; from Web icon it leads users to access at the Mobile Medicine portal for direct download, publication and discussion about mobile medical content. The Local Files Explorer & Indexer module, explores regularly the smart phone PDA memory storage to find new media files, it extracts metadata/descriptors and index them in the local database by the Content Indexer that extract the semantic information from the content and perform the minimal content processing to process descriptors and simple files.
- **search and retrieve complex cross media content**
    - **making queries** on the mobile device searching for locally collected and directly accessible intelligent content. These queries may be full text or advanced taking into account content:

classifications, file naming, grouping for types, taxonomy based, etc., and general semantic descriptors and organizations;

- **browsing on medical taxonomy**, to navigate into the content collection organized for arguments or for intervention type to be performed (for example the pathology), or on the basis of other models and structure, etc.; commonly accepted medical ontologies are accessible;

- **execute complex cross media content** to get support on making decision content on the mobile. Any content is put in execution with a suitable player: PDF player, media player, etc. Cross media intelligent content with internal semantic computing reach the PDA phones as MPEG-21 (for PDA) and XML for i-Phone and thus are played by using their corresponding players. The Mobile Medicine content players may enforce interactivity and intelligence in the content depending on the user profile, user actions, context, GPS, communication with server, etc. i-Phone version provides less behavioural capabilities and no DRM. The descriptor is used for indexing the content and collecting the semantic information. At each execution, a number of use data are collected to be exploited for providing recommendations;
- **receive updates automatically** on the mobiles without user intervention (Download/Update Manager). The module collects content into the local storage of the mobile that is easily larger than 16 Gbyte. It also keeps trace about last date and/or version obtained, goes on the server to verify the presence of new versions, download them and substitute them eliminating the older version;
- **receive personal suggestions** (local recommendations) which are computed on the basis of personal information collected on both server and mobile devices and privately offered on the mobile device: similar content, most used content, suggested content, etc. Similar content takes into account distance among the user profile collected and/or declared and the descriptors of the objects.

Local recommendations are provided off-line and are based on: collected content semantics, user profile, collected use data from executions, navigations and queries. This information is used to make local suggestions such as the presentation of content according to the most played, less played, most recently played, less recently played, alphabetic order, taxonomical order, recently updated, etc. Personal recommendations are computed and provided in real time to the users on PDA and i-Phones. In order to reduce the query time the local database modelling the medical classification structure is provided precompiled into the installable file of the Content Organizers for both PDA and i-Phone.

4.5. Exploiting Mobile Medicine (Metrics)

As anticipated in chapter 3 the Best Practice Network developed provides an engine to monitor data in order to reason upon content descriptors, user profiles, ads profiles, user generated content descriptors, device capabilities descriptors, etc. In

order to provide to the Portal administrator with an effective tool able to identify the general trends of the BPN including downloads and user behaviour, a set of metrics has been defined, he after it is possible to see the results applying these metric on the Mobile Medicine Portal[3]:

Usage of Devices and Platform: as mentioned, the BPN offers a wide range of contents may be accessible by different types of devices (PDA, i-Phone, etc.) and platforms (Linux, Mac Os X, etc.). Figure 4.5 shows the distribution of device/platforms most preferred by users to access content (for both download and direct play). Most downloads are performed via PC-windows; however, also other operative systems (Linux, Mac Os X) and devices are used (PDA, i-Phone, etc.). Therefore, the 12% of users did get access to content by using mobile devices (i-Phone, AxObjectFinder), the 86% by using desktop.

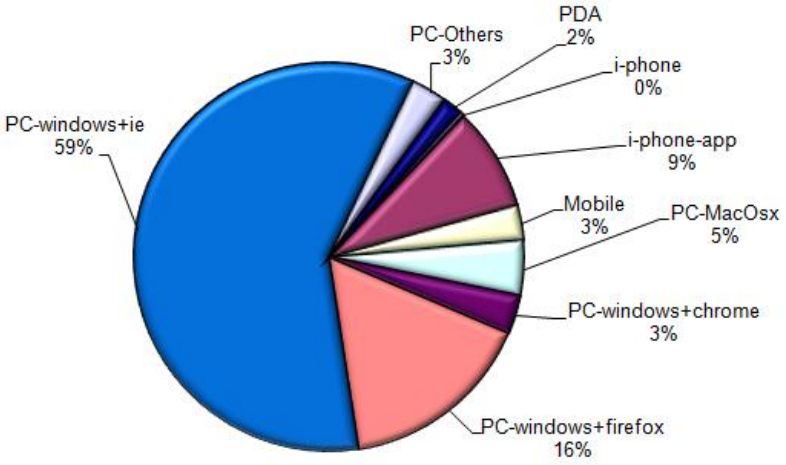

Figure 4.5: Distribution of content access per device/platform.

Proposed and Preferred Content Formats: given the great diversity of content the platform offers, it was necessary to identify which kinds of format are more appreciated by the BPN users with respect to those provided within BPN. In Figure 4.6, the distribution of accessed content format highlights that cross media content is the most accessed format, followed by video and HTML content pages (still in this content, interactive elements similarly cross media).
Moreover Figure 4.7 highlights the strong correlation between the content offered population in terms of formats and the actual distribution of content formats as depicted in Figure 4.5. This graph could estimate what object types are deprecated by users or what classes are more useful for them.

[3] These data have been taken on 2010.

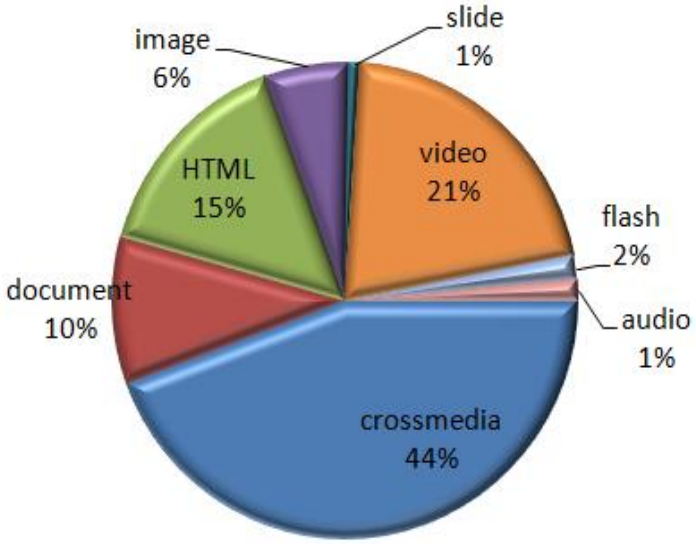

Figure 4.6: Distribution of accessed content formats.

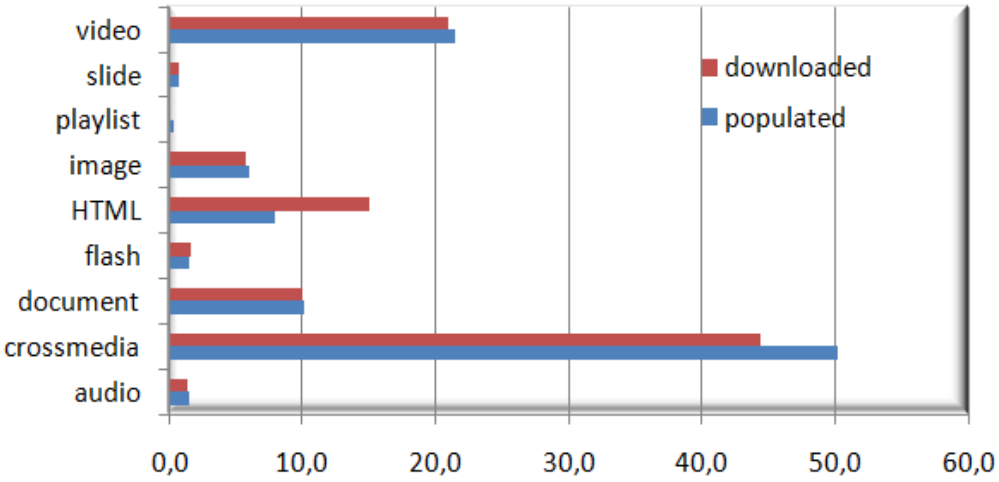

Figure 4.7: Proposed content versus accessed content formats.

Content Access Modality of Users. The content accessed via PC and mobiles may be downloaded or played via the HTTP connection in progressive download. This last modality is possible on PC and i-Phone, while presently it is not viable on Windows Mobile PDA devices endowed with AxObjectFinder tool. Only 1% of content was accessed via AxObjectFinder, while 90% is accessed via online progressive download and only the 10% via download (including the accesses via AxObjectFinder). From this analysis, it is self-evident that users prefer to access the content online rather than to have it on their devices. On i-Phone, it is possible to have online execution of content or download and successive offline play via the Mobile Medicine App. In this case, i-Phone users of Mobile Medicine prefer accessing the videos (38%), HTML (26%) and PDF (21%) while only a small part of them prefer other cross media (please note that dosages and medical calculators are classified as HTML). PDA users prefer accessing the cross media for the 94%. For PDA, cross media include the typical medical tools for dosages, guidelines, checklists, procedures, etc. What is impressive is the adoption of i-Phone for on-line video play. It seems to be the most preferred platform for video play.

Object Stability: In a BPN the distribution of the content accesses is exponential but stable. On such grounds, the Object stability is an interesting metric that allows to understand which is the level of interest for a given content item within the BPN community. This metric measures how many times a content object has been downloaded in a period of time (Figure 4.8a). Moreover, it can be also interesting to normalize the distribution, with respect to the duration of the period of interest from first to last download in the period as reported in (Figure 4.8b).
The comparison of Figure 4.8a and Figure 4.8b gives the evidence of the presence of more explosive content, content used more frequently in less number of days (see Obj-14), while others are more stable such as Obj-1. As to the BPN, it is very important to identify both kinds of object, since the most stable one has to provide a higher quality.

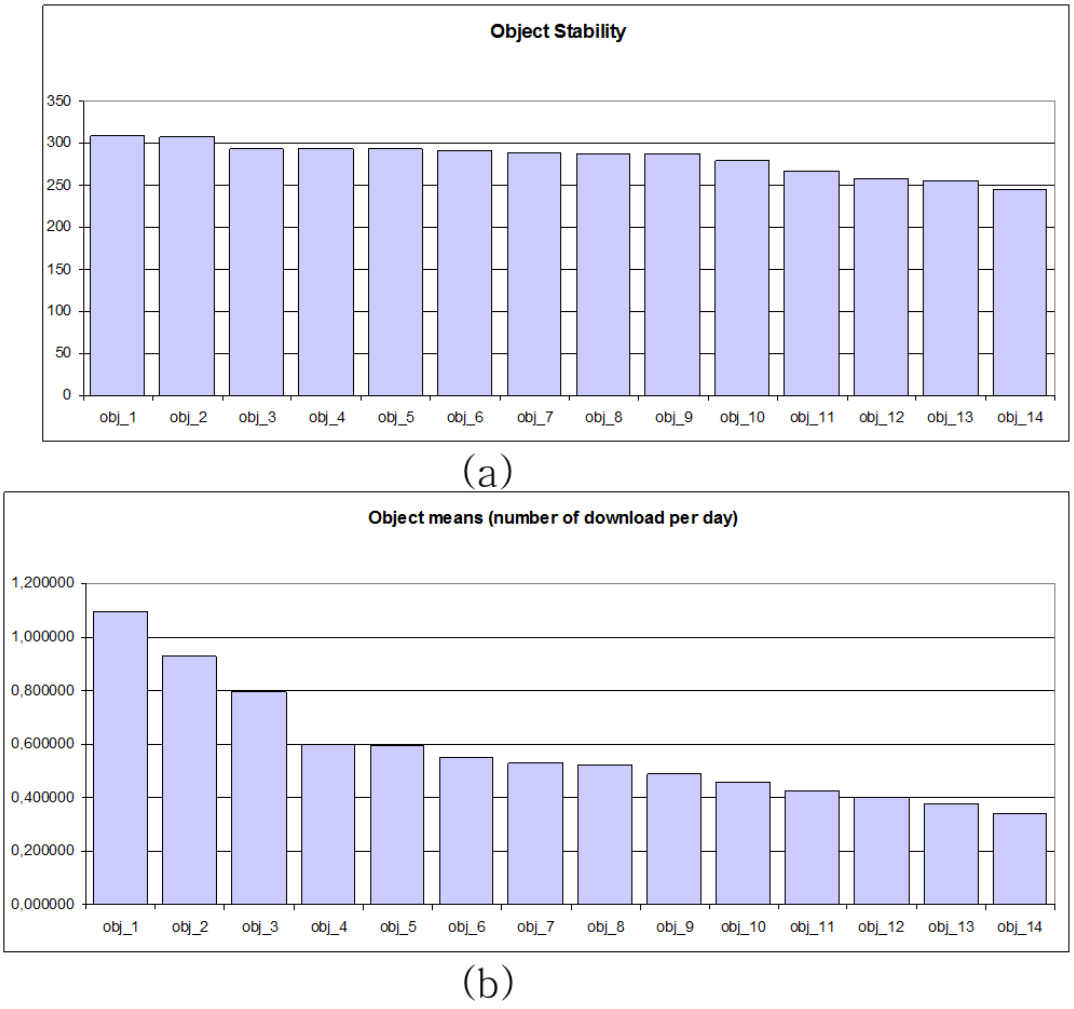

Figure 4.8: Object Stability (a) and mean number of download per day in the stability period (b).

User Fidelity: The aim of User Fidelity measure is to assess in a simple shot which are the most active users in the BPN.
Even in this case, a more accurate measure would lead to go into the distribution of their action over a time period such as that reported in Figure 9, where the distribution in a span of 10 months is presented for a sample of users with different

behaviours (some are returning users, other are sporadic, etc.). It has been also analyzed the user behaviour in terms of number of active days in the period, the so called user fidelity. This simple metric is strongly correlated to the number of content: preferred and uploaded (with a correlation value greater than 75%). In order to assess the user behaviour, it can be useful to see which is the user activity in the sections. In Mobile Medicine, the 63% of users access the portal to get/play a single content even if they return back later and in the next days as depicted in Figure 4.9. Only the 14.3% of them stay for two content items, etc.

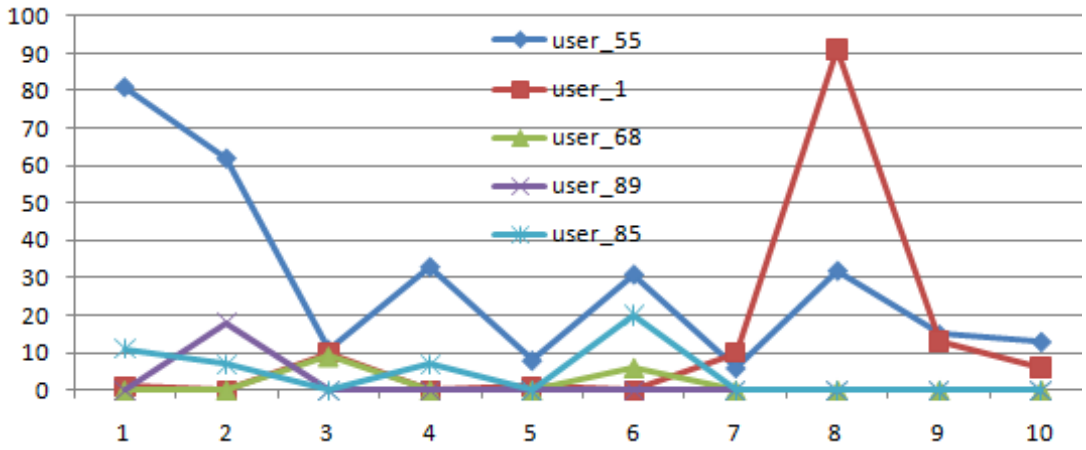

Figure 4.9: User download distribution in the period.

4.6. Metadata and taxonomy

The cross media contents have metadata associated that are the first vehicle to search for objects within the database of the portal. These metadata have to be specified by the users at the upload time of the content and include:

- **Classification in standard Dublin Core**, [DC]: can be produced in more languages by the users or simply by delegating the system to automatic translate them into other languages;
- **Title**: title according to Dublin Core, for each object can be assigned one or more titles;
- **Author**: typically is who has produced the object and not who has loaded it on the portal (the user that makes the upload is maintained in the *user* field (described in the following paragraph) and is automatically identified by the system: in fact only registered users can upload objects);
- **Subject/object**: a concise description of the contents of the object;
- **Description**: description;
- **Associated groups**: one or more groups from the list of possible groups active on the portal to which a content can be associated;
- **Taxonomy**: one or more bindings to medical taxonomy;

- **Visibility:** a content can be accessible in many modalities, for example as a public content (all visitors) or as a private content in one or more groups (only group members);
- **Language of digital content**: this field is useful to specify in which language is the document or the video that has been uploaded on the portal. If the content is a document or a web page containing text, it is parsed and inserted in the search engine based on the language of the document;
- **Metadata Language**.

Other metadata are automatically produced by the system:

- **Format**: technical definition of the type of data (audio, video, image, document, crossmedia);
- **Type:** conceptual definition of the type of data (audio, video, image, document, slide, flash, pdf, html, smil, play lists, etc.);
- **Date** of publication and date of production;
- **User**: is who has uploaded the item on the portal;
- **Full-text indexing**: the text of documents such as: DOC, DOCX, PPT, PPTX, PDF, HTML, etc.

The system is also capable of indexing textual content in multiple languages of resources such as documents txt, doc, docx, pdf and power point.

As previously described, some of the relevant needs in a Medical Platform are related to collecting and indexing digital content and consequently to search and retrieve it in the Platform. In order to do this, classify the content through a specific Taxonomy is fundamental.
The actual Mobile Medicine Taxonomy ([MobMed], [MobMedUM]) has been realised following the next steps:

- State of the Art related to medical platforms:
  - MyOpenCare, [MOC];
  - Esanum, [Esa];
  - etc.
- State of the Art related to the medical libraries and classifications:
  - International Classification of Diseases, [ICD];
  - PubMed, [PM];
  - MeSH Tree Structures (U.S. National Library of Medicine National Institutes of Health, [NLMMesh]);
  - etc.
- Write the ‘ad hoc’ taxonomy as a text, (for details see [MobMedUM]);

- Implement the taxonomy in the Mobile Medicine platform (available both for PC and mobiles).

In order to manage the Taxonomy on the mobile Medicine portal, a set of tools/functionalities are provided:

- **Taxonomy Creation Interface** (Figure 4.10) for editing taxonomy terms. This tool is accessible only to the site administers and it allow them to change the structure of the taxonomy, to add or delete terms;
- **Association of one or more taxonomy term to a content**. This functionality allow all the registered users to associate a content taxonomy (e.g. web page, group, forum, blog, digital content, etc.). This association is usually done at the uploading or creation time (Figure 4.11);
- **Taxonomy translation**: this functionality is automatically managed by the system (20 languages are available);
- **Browsing on medical taxonomy**, to navigate into the content collection organized for arguments or for intervention type to be performed (for example the pathology), or on the basis of other models and structure, etc.; commonly accepted medical ontologies are accessible, (Figure 4.12);
- **Import/export**: e.g. the export functionality can be used to export the taxonomy into RDF/XML.

Figure 4.10: Taxonomy Creation Interface.

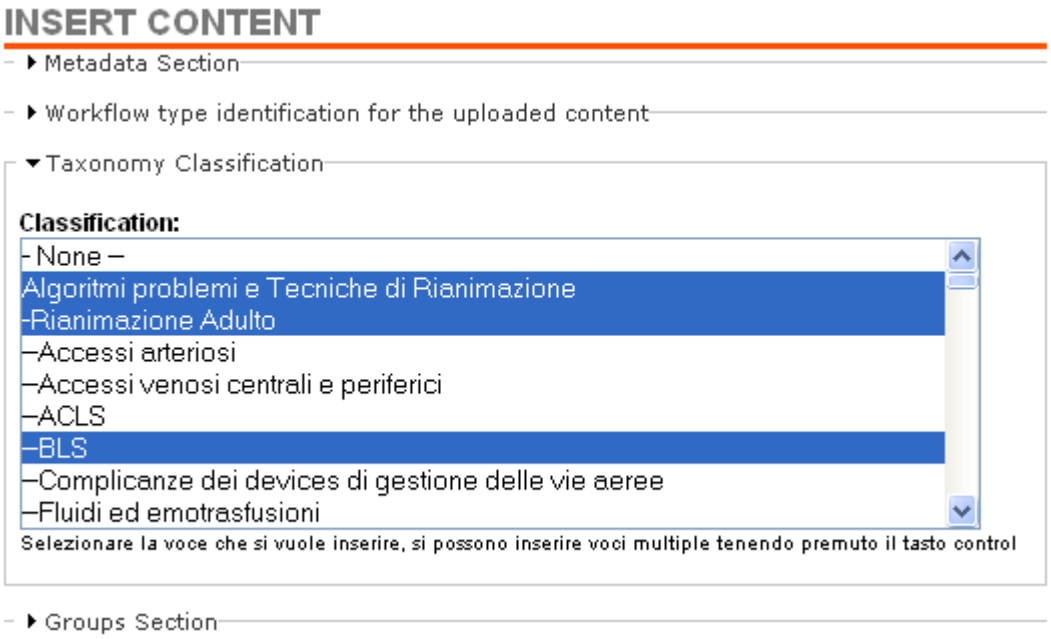

Figure 4.11: Association of one or more taxonomy term to a content.

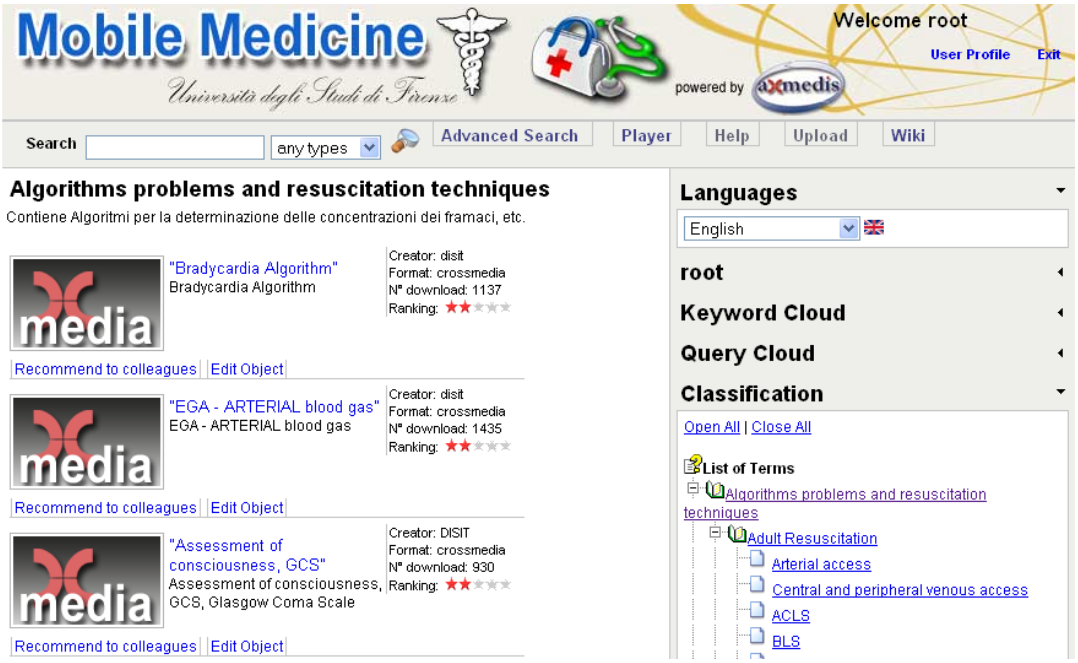

Figure 4.12: Browsing on medical taxonomy.

4.6.1. Dependencies

Internal component	TyType (drupal module, AXCP rule, Java Servlet, Other)	**Notes**
taxonomy	Drupal module (customised)	Drupal core module
taxonomy_context	Drupal module (customised)	Available at [DruTax]

taxonomy_xml	Drupal module	Available at [DruTaxImp]

4.7. Groups

Groups and distribution channels are fundamental in a Best Practice Network, especially relating task of managing user groups and digital contents, provide in a simple manner what users want, make optimum use of the portal's features to produce and manage contents, increase social relationships and knowledge in the domain area, provide a space for discussions, etc. The main features and services are be described in this paragraph, [EC-DE3.1].
It has to be noticed that in the BPN many different user roles are present. Each role is connected to different privileges. Regarding the groups, the following roles, as said in the introduction, are fundamental: visitor, registered user but not group member, group member, group manager, site administrator. As usual each role has specific permissions and the system provides to them different functionalities and services.
The roles of the people within groups are defined because in a BPN it is natural that there are people with major duties and responsibilities (e.g. group manager, site administrator): update the information published and made available to the group members, establish basic behavioural rules, have the technical instruments for possibly take correctional actions, maintain the actual group state (both regarding the content uploaded, the discussion topics and the user behaviour). It is equally fundamental the simple users presence (group member), they are able to play basic actions within the group: the system puts them in the position to carry out requests and to receive services that match their expectations. Finally you must promote the group created within the site: this is why it is essential that some basic services are available even to the simple site visitors and both to all registered users (people not enrolled in the group).

In the BPN many different user roles are present. Each role is connected to different privileges. In order to explain the group services can be useful define the user roles involved in the group and respectively the services provided for them in the portal (the services are not described in details, for more information see [MobMedUM], [EC-DE3.1]):

Unregistered user: people visiting the portal. They can:

- See the portal group list
- Access to 'Public Group Home Pages'

Registered user: people registered to the portal but not yet enrolled to groups. The portal provides to them the same services provided to visitors, moreover they can:

- Request to add a new group

User Enrolled to one or more groups: people registered to the portal and enrolled to one or more groups. The portal provides to them the same services provided to registered user, moreover they can:

- Access to 'private group home page'
- Make use of group forum/ forum topics
- Send a broadcast message to all users enrolled to the same groups
- Invite colleagues/friends to subscribe their groups
- Send specific invitations to users and acquaintances to suggest interesting links
- Directly view the group members list/photo (and easily access to their profile)
- Access to pages/contents related to the group
- Receive notification on group events: via e-mail or personal messages on the portal
- etc.

Group manager: people registered to the portal and managing one or more groups. Note that the ECLAP groups are moderated: each group can have more than one group coordinator that can manage the group. The portal provides to them the same services provided to registered user, moreover they can (see Figure 4.13, Figure 4.14):

- Modify/ maintain group home pages (both private and public)
- Create WEB pages for each group he is coordinator of;
- Cancel a discussion topic in the group's discussion forum he chairs;
- Perform direct registration to the group on behalf of ECLAP BPNET Portal's registered users via their nickname, provided they have already registered to the portal;
- Delete user from the group and/or put them in a black list, block a user from accessing the group;
- Accept or deny group membership requests from portal users;
- Register new group members (from them enrolled to the Platform);
- Promote/remove a member as group administrator.

Site administrator: who has the role of Portal administrator. Regarding the group functionalities, the portal provides to him/her the same services provided to group coordinator, moreover he/she can:

- Create new groups;
- Remove all the group coordinators;
- etc.

Figure 4.13: Group manager view: private home page.

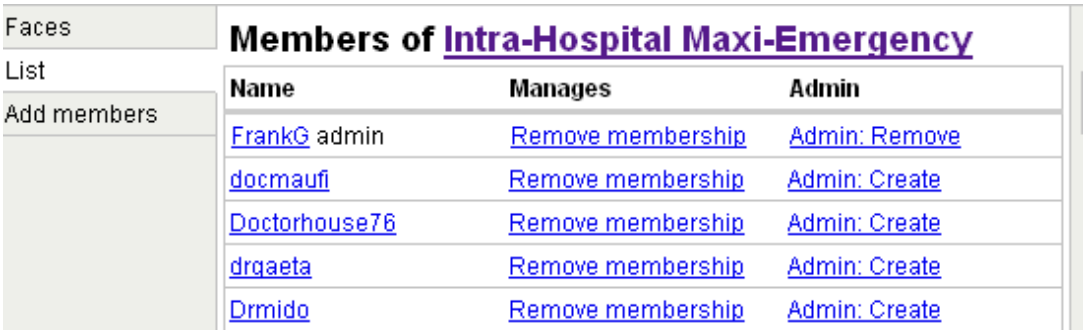

Figure 4.14: Group manager view: members management.

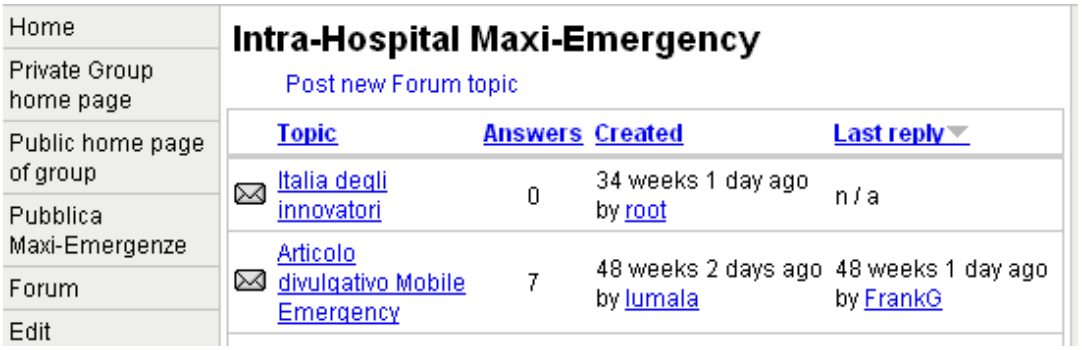

Figure 4.15: Group manager view: forum topics.

4.7.1. Dependencies

Internal component	Type (drupal module, AXCP rule, Java Servlet, Other)	Notes
og	Drupal module (customized)	Available at [DruOg]. This Drupal module contains other additional sub modules needed to improve the functionalities (located in the sub folder 'modules' and reported here after).
og_views	Drupal module (customised)	
og_access	Drupal module (customised)	
og_actions	Drupal module (customised)	
og_forum	Drupal module (customised)	
og_notifications	Drupal module (customised)	
og_panels	Drupal module (customised)	To realise the group home pages (wall)
og_user_roles	Drupal module (customised)	
og_views	Drupal module (customised)	

5. ECLAP – European Collected Library of Artistic Performance

In the last years, the society of information has been permeated by new ways of fruition of multimedia contents on the web. The attention of users is more focused on content oriented web sites, and in most of them several forms of content aggregation are offered in the context of social and best practice networks. Those portals and services are facilitators for content access, while the real content items, the digital essences, are accessible only on the original portal of the content provider because they are referred. This may happen in the digital libraries of IEEE and ACM, which in some cases index metadata and refer other collections via URLs. Among indexing portals, Europeana (i.e., the European Digital Library, [Europeana]) collects cultural heritage metadata coming from several institutions, universities, foundations, museums, schools of art, representing a cultural heritage of the huge European history. Europeana portal collects only metadata, while content files are referred via some URL, [BeNePaSe]. These URLs refer to the original content owner and/or to the Content Aggregator, facilitating the collection. The ECLAP (European Collected Library of Artistic Performance, [Eclap]) ICT PsP project of the European Commission develops a Best Practice Network acting as a Content Aggregator, making use of advanced database and delivery tools for the production and dissemination of the rich multilingual European heritage. ECLAP is establishing a set of best practice guidelines covering performing art key areas, such as metadata and content modelling, mapping metadata standards, semantic enrichment, IPR management tools, business models, ingestion and integration of end-user contributions, education and leisure tools, digital libraries tools. To this end, ECLAP work is going to organize international workshops, conferences, via a number of working groups. The ECLAP also supports clustering between several projects identified by the European Commission: ECLAP places at their disposal all solutions and analyses performed, as well as procedures for transcoding metadata and content, and all the experience and codes to replicate ECLAP solutions. In this chapter, some of the most relevant models and tools studied to realise the ECLAP Best Practice network as a valid Content Aggregator,

is described. Especially those related to the Intellectual Property Rights management on the digital content available in the ECLAP Portal.

5.1. ECLAP major objectives

As described above, the richness and value of the European performing arts heritage is unquestionable. Even though these collections are now being digitized and published online, they remain scattered, and coordination is lacking between digital libraries and the performing arts field; however, there is a high demand for access to this content. ECLAP fills this gap by creating a considerable, and hitherto missing, online archive for all the performing arts in Europe, and providing solutions and tools to help performing arts institutions to enter the digital Europe by building a network of important European performing arts institutions and archives and publishing content collections on Europeana, the European Digital Library, [Eclap]. ECLAP Consortium brings together European leading national performing arts institutions, universities and research institutes. The partners belongs from thirteen countries combine their expertise and scientific minds to achieve the ECLAP goals and the planned activities of the next period (for more details see [EclapParts]).

ECLAP is making use of advanced indexing database and delivery tools for the production and dissemination of the rich multilingual European heritage. This will result in cultural enrichment and promotion of European culture, and in improvements in learning and research in the field of performing arts.

ECLAP is presently distributing more than 63000 content objects, coming from more than 20 prestigious European institutions: images, video, documents, audio, slides, play lists, collections, annotations, etc.

The main European Collected Library of Artistic Performance objectives are:

- bring together Europe's most relevant performing arts content, content never before accessible via the Internet, coming from major institutions; performing art material coming from theatre, dance, music, cinema and film, etc.: representing performances, lessons, master classes, teaching material, etc., in the forms of videos, audio, documents, images, animations, play lists, annotations, interactive content, etc.; available through ECLAP portal and published on Europeana;
- create a stable and open best practice network of European performing art institutions, to help them to exploit digital content and to talk about new technologies and tools;
- providing solutions and services to major performing arts institutions such as: content ingestion, metadata enrichment, content distribution, content aggregation into Europeana, IPR management, content channel visibility, play lists, annotations, multilingual semantic/fuzzy search queries, partner/colleague search, etc.;
- provide solutions and services for a variety of users: teachers, students, performers, researchers, and performing arts lovers for edutainment, infotainment and entertainment.

In order to realise the objectives above described, ECLAP provides the following main functionalities and facilities:

- social service support with registration, discussion groups, mailing and forums, recommendations;
- search, retrieve and play extensive high quality multilingual content;
- access and play a large range of cross media content from video to e-books;
- semantically enrich and contextualize, annotate, aggregate content;
- upload and share multilingual content for professional and User Generated Content, get them indexed with the others;
- comment, annotate, rate and vote on content;
- Register and networking with others users and institutions;
- create discussion groups and forums;
- upload digital resources for professional and User Generated Content;
- wizard tools to solve, manage and define IPR issues and accesses on content;
- distribution and access all content via different devices such as PCs, tablets and smart-phones;
- ingestion and publication of metadata to Europeana, to conquer more visibility at the content;
- usage and exploitation of e-learning facilities for the content providers and in general for registered users;
- statistical analysis related to the actions of the users on the platforms and on the content;
- workshops and events organization.

In this chapter, some of the main features of the ECLAP Best Practice Network are described:

- IPR issues and IPR Wizard;
- content management in groups.

It has to be noticed that all the Best Practice Network features described in the previous chapters can be activated basing on the ECLAP needs.

5.2. ECLAP IPR issues

Performing arts content has a high cultural and economical value. Content that ECLAP partner institutions are going to provide embodies many precious materials.

This material is captured on film, video, audio, images, books, posters, etc.. Some content may be submitted by users of the platform. It has to be stressed that this material has so far been 'locked off' from the circulation of knowledge, never posted on Europeana, and one of the great merits of ECLAP is to make it easily accessible on-line, browse-able and searchable for all, [EC-DE3.1]. In the ECLAP Project content partner institutions are providing content with IPR issues already solved, or that have to be solved during the project: there are existing contracts with the persons and parties involved in the copyrights of the contents to preserve and disseminate these valuables documents (e.g., contracts between institutions and performers, etc., to be able to use their digitalized documents). Moreover content partners may also have an archive in which they store content in addition to the portion that is going to be offered and used for Europeana and ECLAP. ECLAP Project is establishes and provide services to cultural heritage institutions to help them in understanding how their content can be made accessible on internet and to help the performing art collections to respect the current European directives and local/regional laws related to copyright and privacy.

The use of ECLAP Portal and services for performing arts institutions as a multi-channel intelligent content distribution demonstrates how performing arts institutions can benefit from the usage of this powerful platform in conjunction with Europeana for their day-to-day activities. Furthermore, the services that ECLAP provides for collecting, integrating and enriching performing arts content will be the added value by which the institutions will be interested to continue to support ECLAP activities. It needs to be noted that Europeana has provided a core set of interoperable licenses that cover rights information for objects in Europeana [eu_rights]. The ECLAP Platform will be fully compliant with this model: each content uploaded in the ECLAP Portal and sent to Europena will be associated to one of the licenses allowed by Europeana (for details see the following paragraphs), allowing Europeana to access the content/metadata and to inform users about the license associated to each content.

5.2.1. ECLAP General workflow

In order to describe the flow of actions, rules, procedures, etc. that each Content Partner will follow to publish its content on the ECLAP Portal and provide it to Europeana, it is fundamental to introduce the ECLAP overall scenario (Figure 5.1) and some common definitions that is used in the following sections.

Workflow
All the content managed in the ECLAP Best Practice Network (both metadata and digital item files) must follow a precise general workflow (WF) before it can be connected to Europeana via its metadata. Each content has to be:

- uploaded in the ECLAP Portal by a Content Partner4;
- enriched through metadata in the ECLAP Portal (some metadata have to be sent to Europeana and others are necessary to describe and manage the content in the Eclap Portal);

[4] The Content Partner has to accept the ECLAP Terms of Use , [ECToU].

- associated to a set of access permissions (through the IPR Wizard, see next paragraphs);
- associated to an IPR license (also in this case, through the IPR Wizard) to be exposed to the final users.

Content uploaded/ingested are initially available on ECLAP portal with maximum restrictions. While metadata are immediately available for indexing and search for all kind of ECLAP users. Only content presenting a (i) sufficient set of metadata (e.g., Europeana mandatory metadata) and (ii) IPR information and a license defined (one from the set admitted by "europeana:rights"), the metadata will be published on Europeana.

Roles and Users
To clarify the general scenario on IPR related to ECLAP, the Consortium identifies specific user roles that are involved in the IPR management in some manner. They are synthetically described here after:

- **Content Providers (CP).** They exploit the ECLAP services in the future (institutions involved in providing content) and can be: original ECLAP partners or affiliated partners;
- **ECLAP Portal (EP),** the service portal itself;
- **Public ECLAP Users (PU)** of the ECLAP portal: people not registered on the Portal;
- **Registered Users (RU)** on the ECLAP portal: they can be further classified according to their profile (one of the most important fields is their affiliation, Fig. 1). Moreover they have to be subscribed to one or more groups/channels. In addiction, registered users can be:
    - Educational Users (EDU) belonging to a workplace (University, School Research center, etc.), having a specialization (student, researcher, professor, etc.);
    - Not Educational Users (NOT EDU);
- **Europeana**, the service portal itself;
- **Europeana Users (EU)**, public users on the Europeana portal;
- **Trusted Users (TU)**: a registered user (coming from Consortium Partners and Affiliated Partners) with particular privileges, like group responsible, root, etc;
- **WF IPR User**: is a registered user with the privileges to edit the IPR Permission by using the IPR Wizard tool (see next paragraphs);
- **WF Enricher User:** is a registered user with the privileges to edit the metadata connected to a content;
- **WF Validator User:** is a registered user with the privileges to validate all the metadata edited (each validator can choose from a list of languages in which he/she is expert);

- **WF Publisher User:** is a registered user that authorize the publication of the metadata on Europeana.

To be registered on the ECLAP Portal, Public Users have to accept the ECLAP Terms of Use (ToU, [ECToU]) during the account creation as mandatory step to have the access to ECLAP services. The ToU expressively specify that the User is responsible for the correctness and accuracy of the information reported in his user profile (that is considered as a self-certification).

Figure 5.1 shows the general scenario related to IPR involving all the identified roles in the value chain. To simplify the description of the following image, the distinction between the metadata ingestion process and content ingestion process is necessary. Regarding the metadata, the procedure foresees that when the metadata are ingested from the content owner (independently if the ingestion will be done as a single file upload or via a massive ingestion), the metadata will be provided to ECLAP with a CC0 1.0 Universal Public Domain Dedication by default (as defined also in the Europeana Agreeement [eu_agreement]). This license grants ECLAP the rights "to publish, make available, reproduce, distribute, display, transmit, extract, re-utilize and store the Metadata and its derivatives in all forms, formats and media", and "to translate the Metadata (or authorize others to do so) into other languages, create adaptations, summaries, combinations, or extracts of the Metadata". As aggregator of Europeana, ECLAP ingests the metadata and performs also adaptation procedures to make the metadata consistent with the Europeana metadata model. The metadata will be sent to Europeana only when the metadata will be linked to a reachable resource and IPR issues will be correctly defined. Regarding the content ingestion, the procedure is more complex and foresees the following main steps:

- to guarantee the ECLAP services, ECLAP acquires from the content provider a non-exclusive license to use, adapt, distribute, prepare derivative works, display, and perform the Content, in connection with the ECLAP Service (as described in the ToU, [ECToU]);
- Content is ingested applying the stronger restrictions by default (if not in Public Domain), that means that the content is in this phase is not available for the public users, but is only accessible internally for ECLAP Trusted Users;
- Content Provider can relax the access to the content by defining permissions for the final users by using the IPR Wizard tool (if the content is not in Public Domain, see next paragraphs). This operation is performed by WF IPR Users.
- Through the IPR Wizard each Content Provider links a license to the content valid for the ECLAP users (as defined in the Europeana Agreement, [eu_agreement]), basing on the licences that they have on their content;
- Only when all steps are completed the content is published on the ECLAP Portal and the metadata are sent to Europeana.

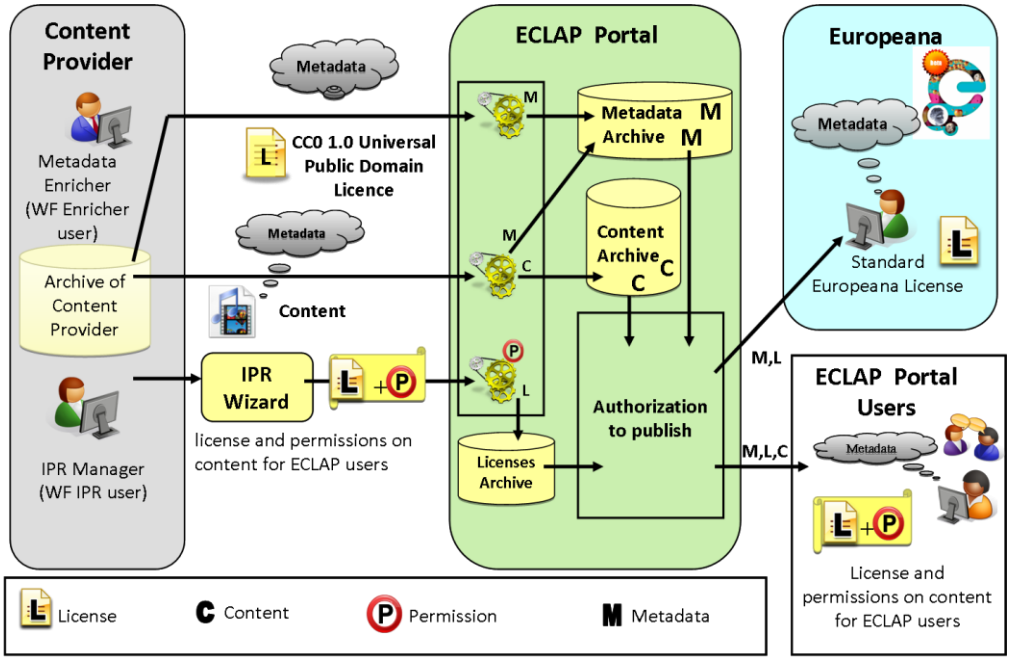

Figure 5.1: ECLAP Overall Scenario.

5.2.2. IPR Issues

As appears clear reading the ECLAP overall scenario, many users with different roles and permission are involved in the ECLAP knowledge workflow. When a Content Partner (in particular a user affiliated to a CP with the "WFIPR user" role) has to select the set of permissions for a specific content, two main aspects have to be considered:

- the relationships among user roles
- the relationship among the permissions

These permissions can be applied on content by the CP by using the IPR Wizard (see paragraph xxx).

Relationships among user roles
The ECLAP users to which the content can be available in the ECLAP Portal are in synthesis (for more details see the previous paragraph). The relationships among them are described in Figure 5.2:

- **Trusted Users (TU):** have all the permissions on the content because they have to manage it in the ECLAP portal;
- **Public ECLAP Users (PU)**: as default they have NO permissions. If an IPR manager assigns a permission to a PU, the system automatically associates the same permissions to all the Registered Users;
- **Registered Users** (RU):
    - **Not Educational Users** (NOT EDU): as default they have NO permissions. If an IPR manager assigns a permission to a PU, the

system automatically associates the same permissions to all the Educational users (EDU);

- **Educational Users** (EDU): as default they have NO permissions. If an IPR manager assigns a permission to a PU, the system doesn't make inference on permissions: only the educational users have the set of permissions on the content.

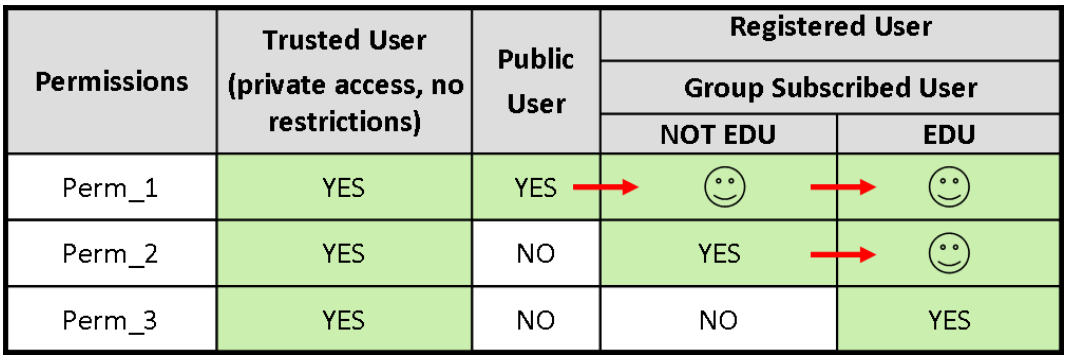

Figure 5.2: IPR Relationships among users.

Relationships among rights/permissions

In The ECLAP Portal many different set of permissions on the content are available. This set of permissions has been realised taking into account both the Best Practice Network Architecture and the ECLAP Content Partners needs.

To realize an questionnaire has been distributed to all Content Partners regarding the following main topics: General contacts; Content and metadata Upload methods; Metadata standards and formats; IPR on content (licenses, permissions, etc.); Collection topics; etc. (for a more detailed description of the ECLAP questionnaire, see Appendix II). The questionnaire results have been necessary to specify the type of permissions they would like to apply to their content. The ECLAP Permissions are classified basing on their type and are (see Figure 5.3):

- **Video:** 8 different permission types;
- **Audio**: 6 different permission types;
- **Pdf/images/html/annotations**/etc: only access permission can be defined (e.g. for a pdf there is no differences between the ‘download’ and ‘play’ permission).

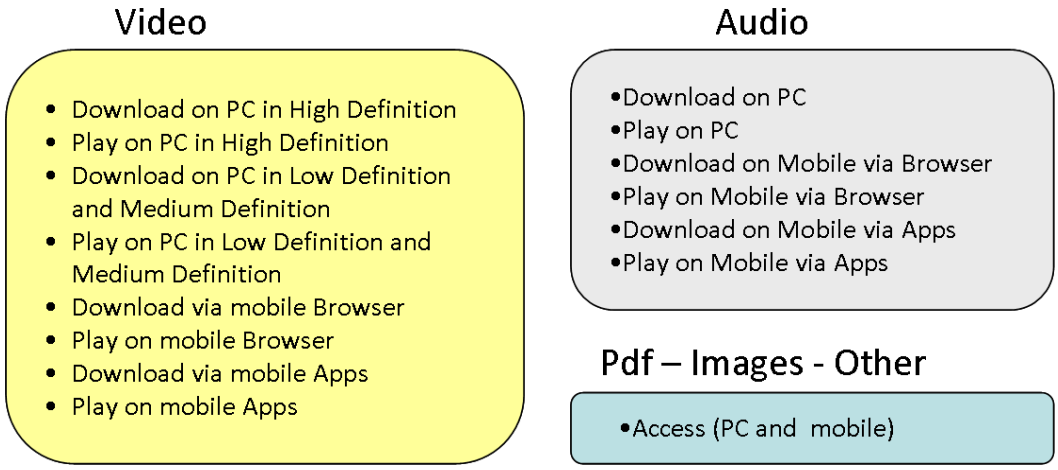

Figure 5.3: ECLAP Permissions.

As it has happened in other studies or in the development of standards (such as: MPEG-21 REL and RDD ontology, IFIP ontology, OMA DRM ODRL rights, CC relationships among rights, etc.) the relationships among the rights identified have been analysed. These relationships have been classified on the basis of the content type to which they are applied. So for the Pdf/images/animations/html/etc. only one permission is present while for the Audio (Figure 5.4) and Video (Figure 5.5) permissions and their relations are more articulated.

The model permissions reported in Figure 5.4 and Figure 5.5are ordered on the basis of two different aspects:

- **Value**, that means that the value of the resource is increasing from bottom to top: for example to assign to a user the permission of download a video/audio gives to the user more values in terms of content than assigning to him/her the permission of only playing it;
- **Control**, that means that the control that can be applied on the resource is increasing from top to bottom: for example to assign to a user the permission of playing a content from mobile through an ECLAP application gives to the Partners more control on user actions (ECLAP applications can recognise the users and register useful information in term of action done etc.) than leaving them playing the content only via a WEB Browser.

These relations have been studied and modelled also basing on logical and technical aspects. They have been accepted by partners and the same model has been used for the first release of the IPR Wizard. In the following figures the arrows are posed to explain that some permissions implicitly involves other permissions.

Here after two samples on audio content (Figure 5.4), useful to describe the permission relations. The relation among permission are represented in as different arrows (same samples can be done also for the video content, Figure 5.5):

- **Unidirectional arrow** between 'Audio download-PC' → 'Audio play-PC': if a Content Partner allows ECLAP registered users to download the audio content, the CP implicitly allows them also to playing it (play via streaming and/or progressive download). This because from a technical point of view, if someone downloads a content (without encryptions or

protection) from the web he/she can play/view it on its PC whenever he wants.

- **Bidirectional arrow** between 'Audio download-PC' ⟷ 'Audio download-mobile-Browser' (note that they are bidirectional): if a Content Partner allows ECLAP registered users to download a content from PC, implicitly allows them also to download the same content from a mobile device. This because the users can download a content via a browser in their PC, then transfer the content into a mobile device, so that the application of a restriction to avoid the download via mobile can be easily moved around and has no sense to be applied. It is also true the vice-versa and, as highlighted in the following figure by the red arrows, also the possibility to play the audio content whenever they want is implicit.

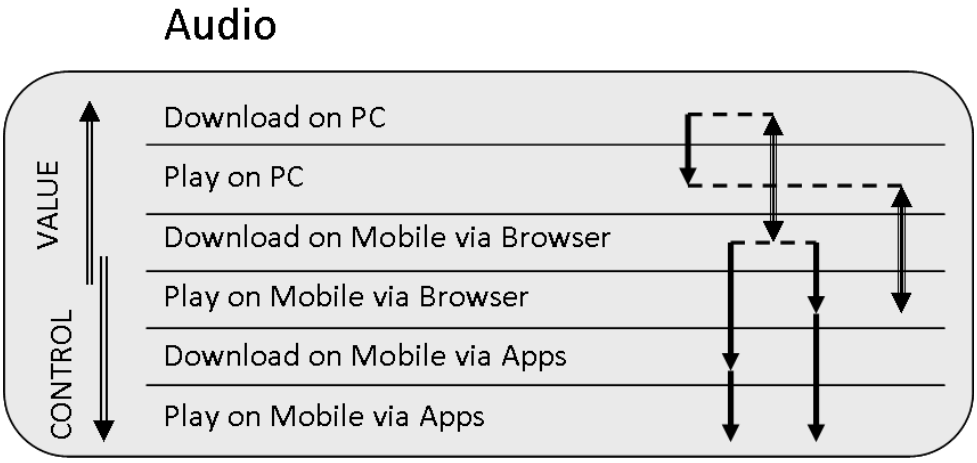

Figure 5.4: IPR permissions relations on Audio content.

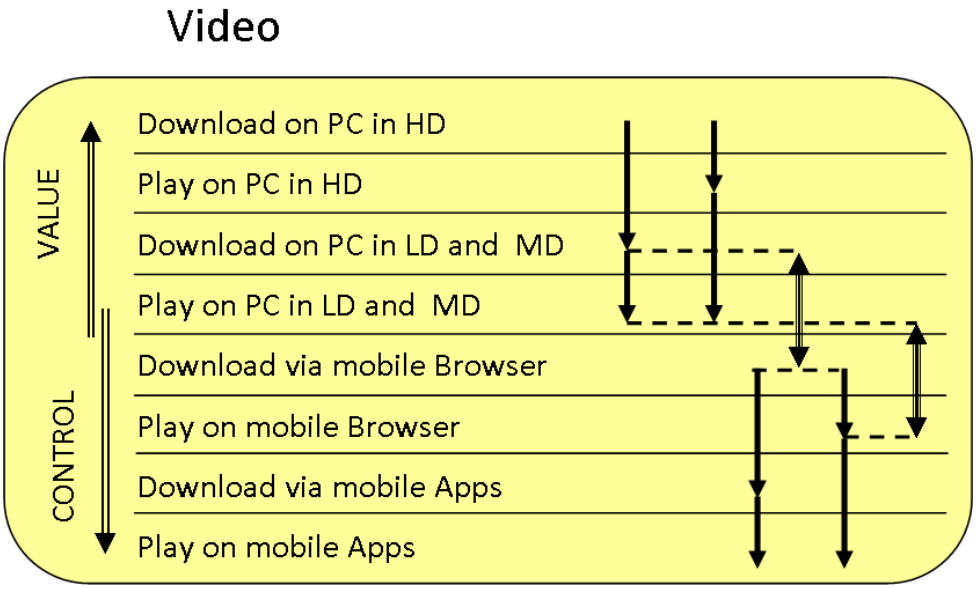

Figure 5.5: IPR permissions relations on Video content.

Sample: Relationships among user and among rights/permissions

All the relations described until now have to be considered when a CP wants to assign a set of permissions to a content: both the relationships among user roles and the relationships among rights/permissions.

In Figure 5.6, it is possible to see an example that considers both these two main features: "If a CP allows all Not Educational users (NOT EDU) to play an audio content from mobile via Browser", the Wizard implicitly assigns[5]:

- basing on the relations among permissions (following the vertical arrows):
  - 'NOT EDU' users can also play the audio content on PC in LD and MD quality;
  - 'NOT EDU' users can also play the audio content on mobile via ECLAP Applications;
- basing on the relations among users (following the horizontal arrows):
  - also 'EDU' users can play the audio content from mobile via Browser;
  - also 'EDU' users can play the audio content on PC in LD and MD quality;
  - also 'EDU' users can also play the audio content on mobile via ECLAP Applications.

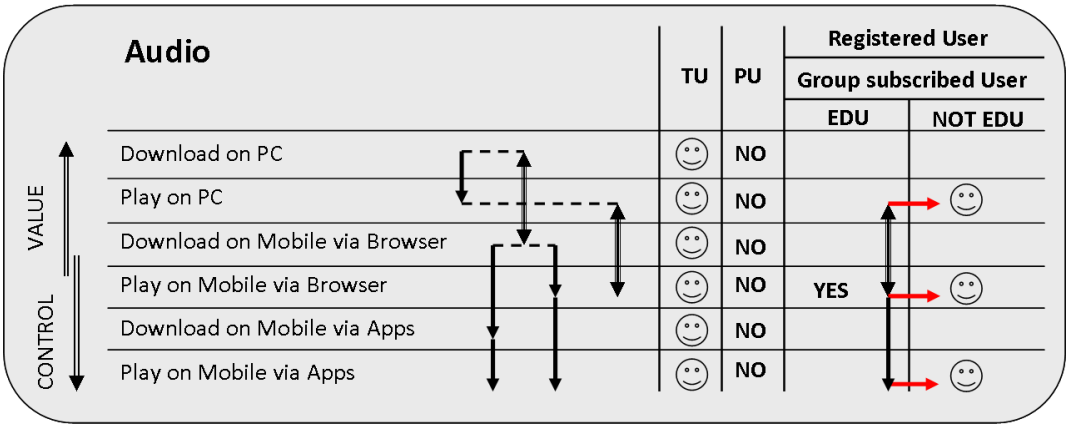

Figure 5.6: IPR permissions relations and relationship among users on Audio content.

After the permissions definition, the CP has to select one license to be associated to the content. This license is compliant with what has been defined by Europeana and reported in the 'Guidelines for the europeana:rights metadata element', [eu_rights].

Moreover, it has to be remembered that another relevant issue to be solved is to assign a licence to each content. The IPR Wizard proposes to the CP to choose a license among the set of licenses allowed by Europeana and on the basis of the permissions defined on the content uploaded. This means for example that if a CP define some restrictions on its content (e.g. 'this pdf can be accessible only by students'), the IPR Wizard does not allow the CP to associate to it the 'Public Domain Mark'.

[5] Because of the complex structure of the permissions relation, it is more simple describe this sample starting from the 'relations among permissions'.

If the CP has classified the content as Public in ECLAP (all permissions allowed to all ECLAP users, that also means no restrictions), the IPR Wizard creates a 'Public Model' and proposes at the CP to choose one of the following licenses:

- Public Domain Mark (http://creativecommons.org/publicdomain/mark/1.0/)
- CC – Zero (universal, http://creativecommons.org/publicdomain/zero/1.0/)
- Rights Reserved - Free Access (http://www.europeana.eu/rights/rr-f/)

While if the CP has imposed some restrictions on the content (has allowed only some permissions or only some users), the IPR Wizard creates a 'Not Public Model' and proposes at the CP to choose one of the following licenses:

- CC BY (v3.0 Unported, http://creativecommons.org/licenses/by/3.0/)
- CC BY-SA
- …
- CC BY-NC-ND
- Rights Reserved - Paid Access (http://www.europeana.eu/rights/rr-p/)
- Rights Reserved - Restricted Access (http://www.europeana.eu/rights/rr-r/)
- Unknown (http://www.europeana.eu/rights/unknown/)

When accessing to a content ECLAP users have to respect the license associated to that content. This license is visible through a link in the ECLAP Portal (and the url of the license will be sent to Europeana as one of the mandatory metadata related to the content - 'europeana: rights').

5.2.3. IPR Wizard

The role of the IPR Wizard is relevant to help the Content Partners to assign the right permissions and to choose the license (one of that allowed by Europeana, [Eu_rigths]) that is associated to it when it is published on the ECLAP Portal and in Europeana. The association of the content to a license and the definition of the permissions are a mandatory conditions for publishing the content. In order to satisfy this conditions, the IPR Wizard performs to the CP a set of questions and produces an **IPR Model**. This IPR Model can than be assigned to a content or to a set of contents.

The Wizard performs questions to the CP to assign:

- assign the permissions on content access on the ECLAP Portal (e.g. content public, content accessible only to DSI educational users, etc. only to educational users, etc.);

- a license to a content (mandatory for EUROPEANA);
- a provider licensing page (an ECLAP Portal page in which you can add licensing information).

When a CP will use the IPR Wizard to assign permissions and the license to a specific content, the IPR Wizard knows the following information:

- type of content (audio, video, document, image, animation, etc.), object Identifier (AXOID), the user and the Content Partner that is using the Wizard, groups/channels to which the content will be visible, etc.;
- relationships among user roles;
- relationships among rights/permissions.

Moreover it has to be noticed that not all the CP users can modify the IPR on the content. Only a Registered User having the following features can do that. A user that:

- has sets his/her affiliation on his/her User Profile (specifying to which Content Partner he/she belongs, e.g. 'DSI');
- is subscribed to a group and is also an IPR manager of his/her group6 (e.g. 'DSI IPR Manager').

Obviously these users can manage the IPR only on their contents, or rather only on only uploaded in the portal from people with the same affiliation (e.g. content uploaded by DSI affiliated users).

IPR Model

An IPR Model contains, (see Figure 5.7):

- Model Details: IPR Model name, description, etc.;
- a set of permissions;
- a licence (Creative Commons, etc.);
- a Publisher ECLAP page (related to one of the Content Partners);
- an IPR ingestion identifier (needed to assign the IPR Model to the contents).

[6] It has to be noticed that Each Content Partner has its group (and only one) on the ECLAP Portal.

Figure 5.7: IPR Model (ECLAP interface).

The IPR Model starts with "All permissions for TU and no permissions for the other users" and a logic has been designed to manage the relations among permissions and among the kind of users. In Figure 5.8 the example on "Audio content", described in the previous paragraph, is visible: what happens "If a CP allows all Not Educational users (Group User) to play an audio content from mobile via Browser". Appears clear that as the Wizard implicitly assigns[7]:

- basing on the relations among permissions (following the vertical arrows):
  - a 'Group User' can also play the audio content on PC in LD and MD quality;
  - a 'Group User' can also play the audio content on mobile via ECLAP Applications;
- basing on the relations among users (following the horizontal arrows):
  - also a 'Group and Educational User' can play the audio content from mobile via Browser;
  - also a 'Group and Educational User' can play the audio content on PC in LD and MD quality;
  - also a 'Group and Educational User' can play the audio content on mobile via ECLAP Applications.

In this case not all permissions to all users are allowed so The Creative Commons Licences can not be associated this IPR Model, so the user can choose the licence from one of the restricted licences allowed by Europeana, Figure 5.8:

- Rights Reserved - Paid Access (http://www.europeana.eu/rights/rr-p/);

[7] In this case the IPR Wizard Interface available on the ECLAP Best Practice Network [Eclap] is shown.

- Rights Reserved - Restricted Access (http://www.europeana.eu/rights/rr-r/);
- Unknown (http://www.europeana.eu/rights/unknown/).

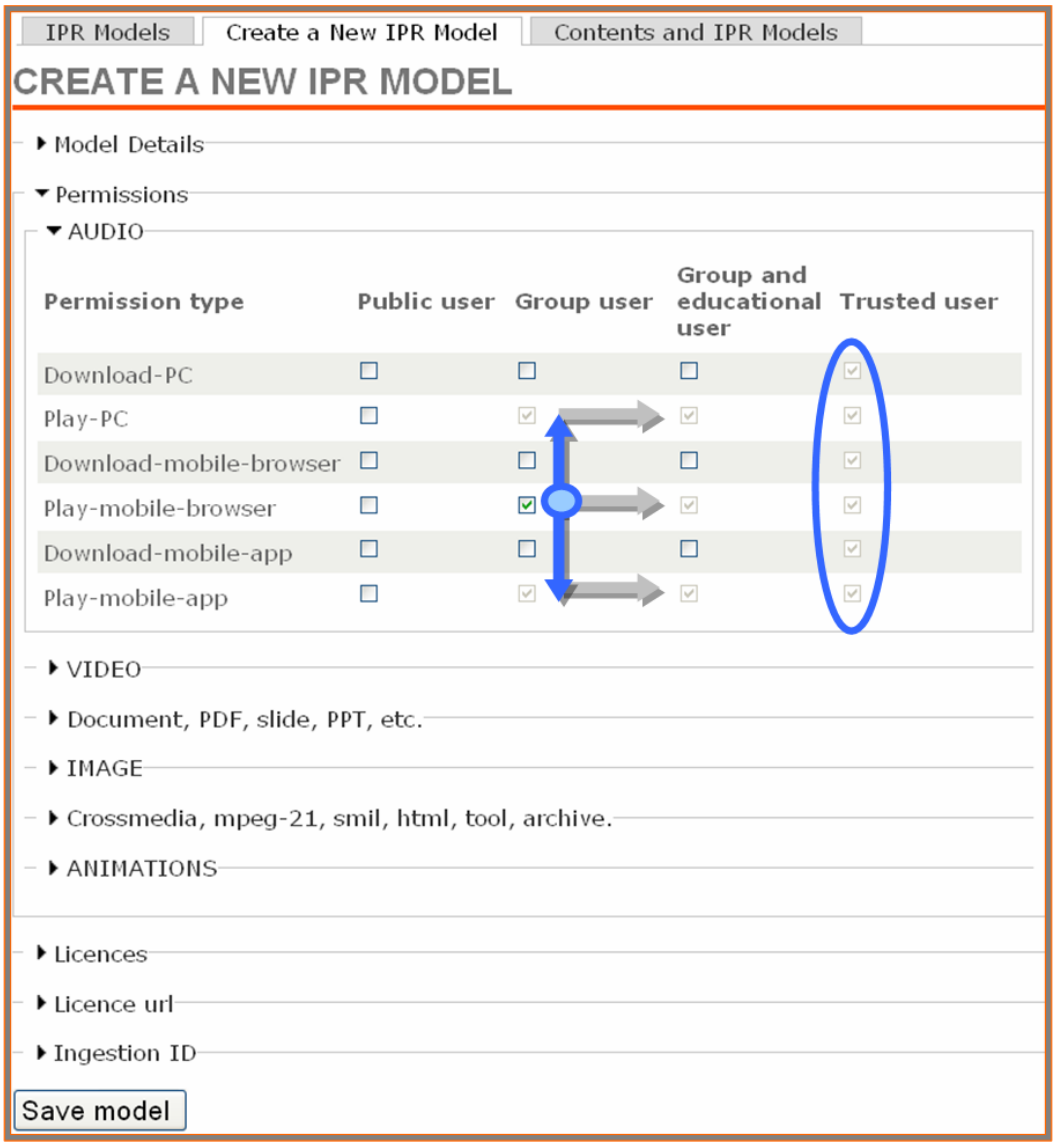

Figure 5.8: relation among permissions on IPR Model (ECLAP interface).

Figure 5.9: Restricted licences on IPR Model (ECLAP interface).

While in the case in which a CP creates an IPR Model in which all the permission are allowed to all the Users, it possible to choose the licence from one of the

Creative Commons ones, Figure 5.10. In both cases a link to the licence is present so the CP can read it and verify the rights.

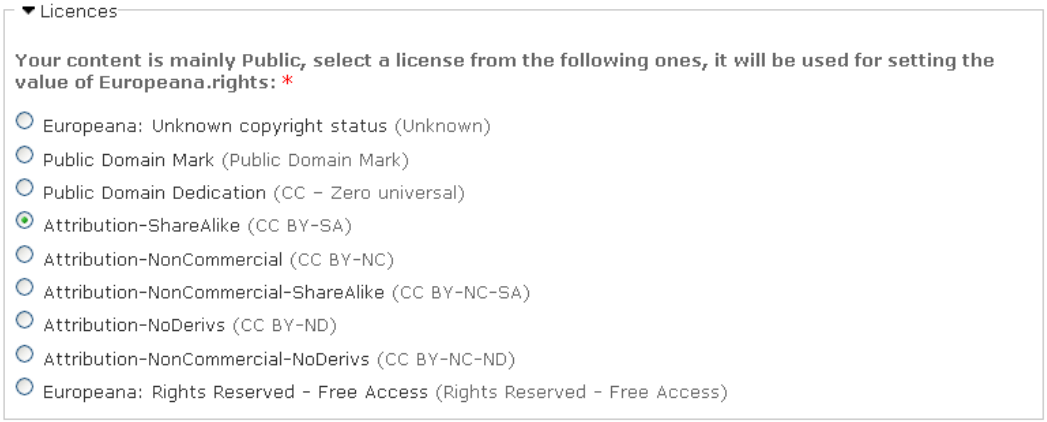

Figure 5.10: Public licences on IPR Model (ECLAP interface).

5.2.4. IPR Wizard Data Model

To store IPR information the tables ipr_license and ipr_permission are used. Table ipr_license contains the association of the license url identifying the license type (e.g. http://creativecommons.org/licenses/by/3.0/) with a human readable description. The table fields are:

- licenseTypeUrl, VARCHAR(255);
- description VARCHAR(255).

Table ipr_permission will store the association of each content axoid with its permissions. The table fields are:

- axoid, VARCHAR(60);
- permission, VARCHAR(30);
- userType, ENUM (‘trusted’, ‘public’, ‘registered’, ‘edu’, ‘notEdu’).

The permission values are:

- download-pc
- play-pc
- download-pc-hd
- play-pc-hd
- download-pc-ldmd
- play-pc-ldmd
- download-mobile-browser

- play-mobile-browser
- download-mobile-app
- play-mobile-app
- access

5.2.5. IPR Wizard Dependencies

Internal component	Type (drupal module, AXCP rule, Java Servlet, Other)	Notes
iprWizard	Drupal module	Custom module (the rules to manage the relations among the permissions are in JavaScript while the other features in PHP)
axmedis	Drupal module	Custom module
eclapworkflow	Drupal module	Custom module
workflow	Drupal module	Drupal core module
node	Drupal module	Drupal core module

5.3. Content management in groups

Given the large amount of contents uploaded on the portal, the development of a service that would allow the group managers to manage a set of contents (and not one by one) was deemed essential, [EC-DE3.3.1]. The Content Management service developed, offers three main functionalities that are detailed here after (see):

Before proceeding in the description of individual functionality is useful to recall some present relations between users, groups, and affiliated partners to the project:

- many groups are available in the platform: groups for discussions (create basing on the same interests), technical groups (for the site maintenance), groups associated to ECLAP affiliated Partners;
- each affiliated Partner has its group in the platform;
- each user has to compile his/her personal profile in the platform specifying the affiliation (one affiliation for each user);
- a content is uploaded by a user or by an affiliated partner (one content is automatically associated to one publisher: the affiliated partner);
- each content can be uploaded by only one affiliated partner but can be associated to more than one partner (mean that will be retrievable in more than one group object list).

Moreover each object in ECLAP has (at least) the following features:

- has a title of the object;
- **Hidden**. An object can be hidden/not hidden. An obj is hidden when it is not indexed: it will not appear in the lists of objects (e.g. Last Posted, Featured, Popular etc.);
- **WF type**. An object can be assigned to different workflow types. Now the WorkFlow (WF) available type are: Internal, ECLAP, Test, Europeana;
- **Void content**. An object can be void/not void. An object is void if only the metadata has been uploaded in the ECLAP Portal without the real content (e.g. in the portal are present only metadata but not the image or the pdf/video, etc.);
- **Published in Europeana**. An object can be just published in Europeana or NOT (this means that the metadata are all present, corrected and validated so that they can be sent to Europeana);
- **IPR Model applied**, the IPR Model associated to the content;
- **IPR Model type**. A model can be public or not public depending on the permissions and on the license chosen to access to it;
- **Published by**. The Affiliated Partner (or Content Partner) from which the content has been updated;
- **WF state**. A specific workflow has been designed in order to better organise and manage the operations that have to be done on the objects, before classify them as collected by Europeana. So a set of WF states have been created (Uploaded, Under-Enrichment, Under-IPR, Under-AXCP, Under-Validation, Under-Approval, Published, [EC-DE3.1]).

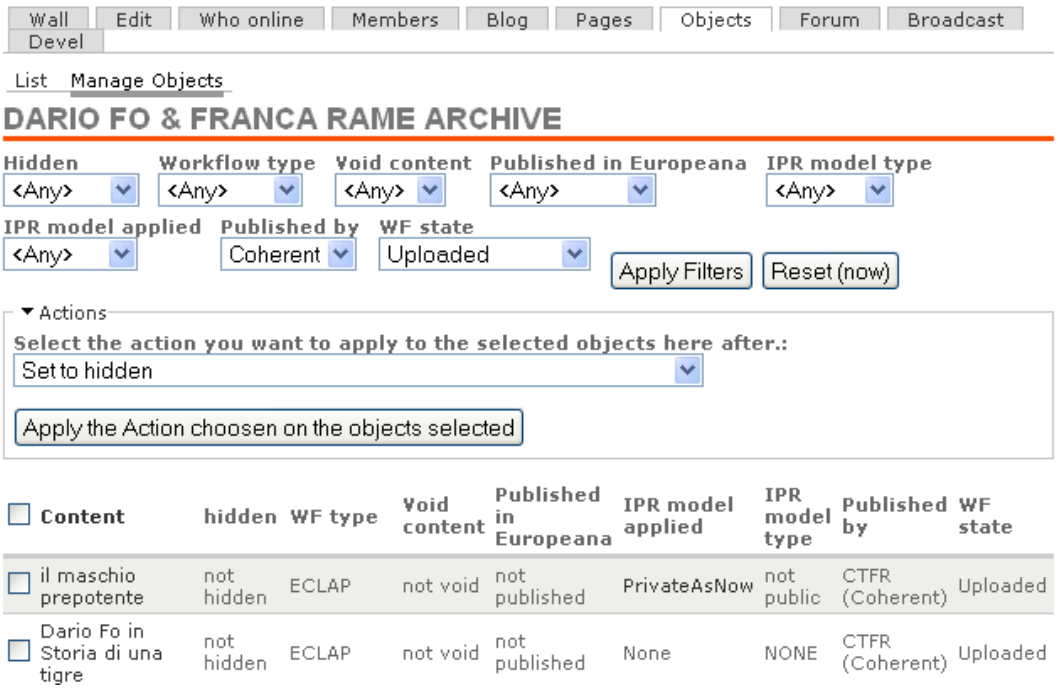

Figure 5.11: Content management in group (ECLAP interface).

5.3.1. View the contents

As visible in Fig.5.4, when a Group Manager goes to the 'Manage Objects' page, can access to the list of contents associated to his/her group. The list of objects is organized as a table in which the rows represent items in columns and lists the characteristics deemed the characteristics deemed most important for content management (classification, search, update content, etc.):

- **Content.** Is the title of the object;
- **Hidden**. (indexed in the BPN or not);
- **WF type**. Workflow type related to each object (it can be one of the following values: Internal, ECLAP, Test or Europeana);
- **Void content**. An object can be void/not void (void means that only the metadata are present in the ECLAP portal);
- **Published in Europeana**. An object can be just published in Europeana or not;
- **IPR Model applied**. To each object can be applied an IPR Model;
- **IPR model type**. An IPR Model can be Public or Not Public, as described above (it depends on the permissions chosen on the content by the CP);
- **Published by.** The Affiliated Partner (or Content Partner) from which the content has been updated;
- **WF state**. The workflow state in which the object is (Uploaded, Under-Enrichment, Under-IPR, Under-AXCP, Under-Validation, Under-Approval or Published).

5.3.2. Filter the contents

As visible in Figure 5.11, when a Group Manager goes to the 'Manage Objects' page, can access to the list of contents associated to his/her group and also filter the list. The filter is really useful in order to selected set of objects having the same features. The Group Manager can use one or more filters in order to have detailed lists of objects. The Group Managers can filter the objects on the basis of following features:

- **Hidden.** (indexed or not);
- **WF type**. Workflow type related to each object (Internal, ECLAP, Test or Europeana);
- **Void content**. An object can be void/not void;
- **Published in Europeana**. An object can be just published in Europeana or not;
- **IPR Model applied**. To each object can be applied an IPR;
- **IPR model type**. An IPR Model can be public or not public;

- **Published by**. The Affiliated Partner from which the content has been updated;
- **WF state**. The workflow state in which the object is (Uploaded, Under-Enrichment, Under-IPR, Under-AXCP, Under-Validation, Under-Approval or Published).

5.3.3. Make actions on the content

Each Group Manager can select a list of objects and apply to the content selected an action. A list of actions is present in the platform, not all the actions are available for all the Group Managers, Figure 5.11.
As declared in the previous DE the platform is structured in way in which the users are organised/classified assigning them one (or more than one) role, to each role correspond a set of permissions and action that can be done in the platform. The Group Manager is one role and have some permission related to actions that can be done in the groups to organise the member, moderate the discussions, access to the 'Manage Objects' page, etc. Other fundamental roles (in addition to the "Group Manager") that are relevant to go in details in this description are: 'Publisher', "IPR manager" (regarding a specific Affiliated Partner), 'site administer'.

- Actions available for all the GMs:
    - Set to hidden (e.g.: set all the objects selected in the check list to 'hidden');
    - Set to not hidden;
- Actions available for all the GMs that also have the 'Publisher' role:
    - All the previous actions;
    - Publish to Europeana;
    - Set the WF type to Internal;
    - Set the WF type to ECLAP;
    - Set the WF type to Europeana;
    - Set the WF type to Test;
- Actions available for the site administer on all the groups of the platform:
    - All the previous actions;
    - Delete object;
    - Copy metadata from DB to object;
    - Copy metadata from DB to object;
    - raw resource and xml extraction and video production (i-Phone, LD, MD, HD);
    - video production (LD, MD, HD).

Moreover an additional action will be provided (now is not available), to the Group Managers that also have the "IPR manager" role of the objects affiliated to their group, is the following one: "Apply the same IPR model".

5.3.4. Dependencies

Internal component	Type (drupal module, AXCP rule, Java Servlet, Other)	Notes
iprWizard	Drupal module	Custom module
axmedis	Drupal module	Custom module
eclapworkflow	Drupal module	Custom module
workflow	Drupal module	Drupal core module
node	Drupal module	Drupal core module
Copy metadata from DB to object	AXCP rule	JavaScript
Copy metadata from DB to object	AXCP rule	JavaScript
raw resource and xml extraction and video production (i-Phone, LD, MD, HD)	AXCP rule	JavaScript
video production (LD,MD,HD)	AXCP rule	JavaScript

6. APRE Toscana

The Sportello APRE Toscana [APRETos], promotes Italian and Tuscan participation in the European Union's Programmes of research, development an Regional APRE Committee – a network of public research institutions, associations and private enterprises – supports, cooperates and provides the strategic policy for APRE's development activities. The Sportello APRE Toscana organises training events and, produces and disseminates documentation, and provides assistance to research, innovation and production professionals.

6.1. APRE Toscana functionalities

To understand and disseminate correct and updated information on EU programmes APRE Toscana refers to national APRE Nazionale [APRE], and CORDIS (Community Research and Development Information Service for Science, Research and Development, [Cordis]) which is the main source of information on the framework programmes, research and innovation under the structural funds and related programs.

APRE Toscana provides an information service on call active and constant on deadlines, as well as on news, meetings, conferences and activities related to the framework programme and various opportunities in collaboration with *APRE Nazionale* and with professionals. APRE Toscana provides this service through its website, a periodic newsletter profiled on the needs of those who register at the site, online discussion groups and an Office open to the public.

APRE Toscana main objectives and services offered:

- **Information**: the portal is structured as a blog. This form of communication (see chapter 2) has been considered the most useful to provide clear and direct information to the users, Figure 6.1;
- **Assistance/Promotion**:
    - Promote and support Tuscany and Italian research through a real support to the design;
    - Offer personalised services for Tuscan enterprises, research institutes and universities (customized information on the Framework Programme and related opportunities, partner search,

consulting, seminars and training courses, personalized services, etc.);

- Generate greater visibility at international level;

- **Cooperation**:
  - Boot the collaboration of research centres and Tuscan enterprises;
  - Increase the number of projects attracted, developed and managed by Tuscan partners;

- **Enhancement**:
  - Enhance valorisation of research results, especially with regard to the cultural and productive issues strategic for Tuscany.

Figure 6.1: Sportello APRE Toscana.

In this chapter the main features of the Best Practice Network APRE Toscana are described:

- Groups
- Newsletter

- Notifications
- Invites
- Contacts ingestion

It has to be noticed that all the Best Practice Network features described in the previous Chapters can be activated basing on the APRE Toscana needs.

6.2. Groups

In APRE Toscana Portal any registered user can create a group. Requests for activation of a group are parsed by the Coordinator of APRE Toscana and by the APRE Toscana regional Committee. The groups are usually dedicated to the dissemination of content and discussion regarding some thematic aspects, or the creation of a thematic channel.

If a APRE Toscana Registered user is interested in creating a new group or wants to become a moderator of a group (just active in APRE Toscana) has to send an email to the site administer, available at [APRE] in which he/she indicates:

- a title for the Group;
- one or more persons registered to the portal (username) wishing to participate;
- proposals and objectives of the Group;
- type of content that could be loaded into group
- type of user that the group might attract

Figure 6.2 is a screenshot of the wall (or home page) of the “Development” group: it is dedicated to the development and maintenance of the APRE Toscana portal. The participation in the group is restricted to members of the APRE Committee. Group participants define the strategies and improvements to be made to the portal in order to improve the service. An e-mail to contact the development team is provided.

Figure 6.2: Group wall (APRE Toscana interface).

The main technical functionalities of the BPN groups have been just described in the chapter related to the Mobile Medicine Project (see chapter 4).

6.3. Newsletter

APRE Toscana is mainly focused on assistance, promotion, cooperation, enhancing the valorisation of research results. In order to realise all these objectives a system to: manage messages and communications among users, provides a newsletter, enhancing personalised notifications is fundamental.
The main way that the APRE Toscana users have to receive information is the Newsletter (Figure 6.3). The system automatically send information related to the portal actions:

- events organised
- digital content uploaded
- changes on pages/blog/forum/groups/ etc.
- etc.

Figure 6.3: APRE Toscana Newsletter.

In order to manage what types of information has to be sent and with which frequency, each user can mange it going on his/her User Profile on the APRE Toscana portal, (for more details see next paragraph on Notifications).

6.3.1. Dependencies

Internal component	Type (drupal module, AXCP rule, Java Servlet, Other)	Notes
mimemail	Drupal module	Available at [DruMime]

6.4. Notifications

The Notification services are accessible from the User Profile page (Figure 6.4), each user can subscribe to a notification type and the system automatically send him/her notifications through the Newsletter (previously described).

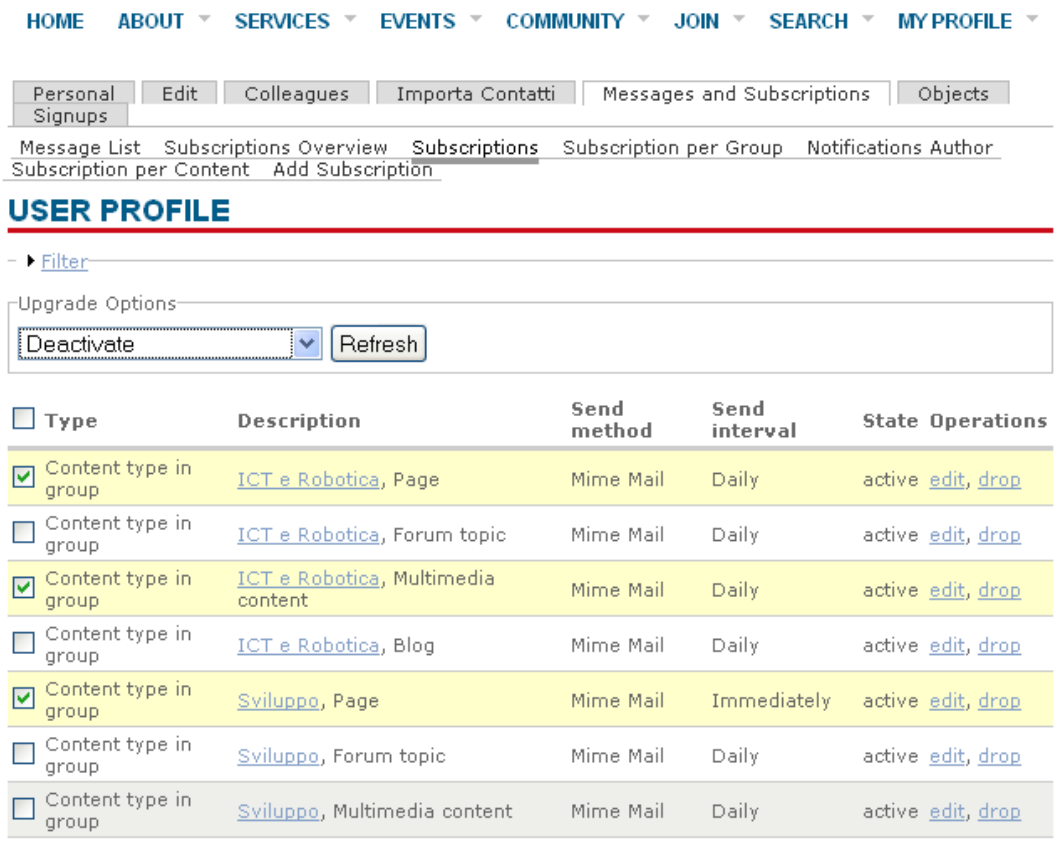

Figure 6.4: Notification services (APRE Toscana interface).

In order to control the Notifications, each user can:

- Subscribe to an event of the portal. It means receive notifications when: a page is updated, a news is published on the blog, a new group is made, etc.;
- see all his/her actual subscriptions and deactivate some of them: this is what is represented in Fig. 4 (the user has selected some notification and wants no notification on them);
- choose the method that the system has to apply to send notifications (e-mail, messages on PC, etc.);
- temporary disable subscriptions;
- etc.

6.4.1. Dependencies

Internal component	Type (drupal module, AXCP rule, Java Servlet, Other)	Notes
notifications	Drupal module	Avalable at: [DruNot]
og_notifications	Drupal module	Avalable at: [DruOgNot]

6.5. Invites

Another way to Promote the APRE Toscana initiative and to increase the number of projects and people that can be involved and interested in the project is to invite user to visit the portal.
The Portal provides the Invite service, each registered user can send to people an automatic message sent by the platform only giving the new user e-mail, Figure 6.5.

Figure 6.5: Invites (APRE Toscana interface).

6.5.1.Dependencies

Internal component	Type (drupal module, AXCP rule, Java Servlet, Other)	Notes
invite	Drupal module	Available at: [DruInv]

6.6. Contacts Ingestion

The main functionalities are present and describe in the following paragraphs: Import contacts from e-mail providers and Social Networks, Import contacts from file, Import contacts view. These functionalities have been developed on order to import contacts that each user has previously registered in other modalities, than manage them in order to invite new users in the Portal or to see if one of the contacts is just present in the Portal.

The main functionalities provided by the platform are:

- Import of contacts from e-mail providers and Social Networks (SN): in this case the user has to select the e-mail provider or the SN from the provided list (Figure 6.6) and fill in the requested fields;
- Import of contacts from file: each user has to select the file in which are the contacts to be imported, Figure 6.7;
- Visualisation of all the imported contacts, Figure 6.8. The contacts are visible in a table format in which each row represents a contact and the columns the most relevant details related to the contact.

Imported contacts table columns, Figure 6.8:

- e-mail
- name
- surname
- nickname (present if enrolled in the platform)
- state of each contact in the platform:
    - *present*. The contact is registered to the Platform;
    - *invited*. The contact is not registered to the Platform and the user has sent to him/her an invitation via e-mail;
    - *connected*. (the contact has made request to be enrolled in the site and, if it is provided in the settings of the platform, is awaiting for the approval from the site administrator.);
    - *new*. The contact is not registered to the Platform;
    - *new_nomail*. The contact is unknown and is classified as not registered in the platform (no e-mail is not available).

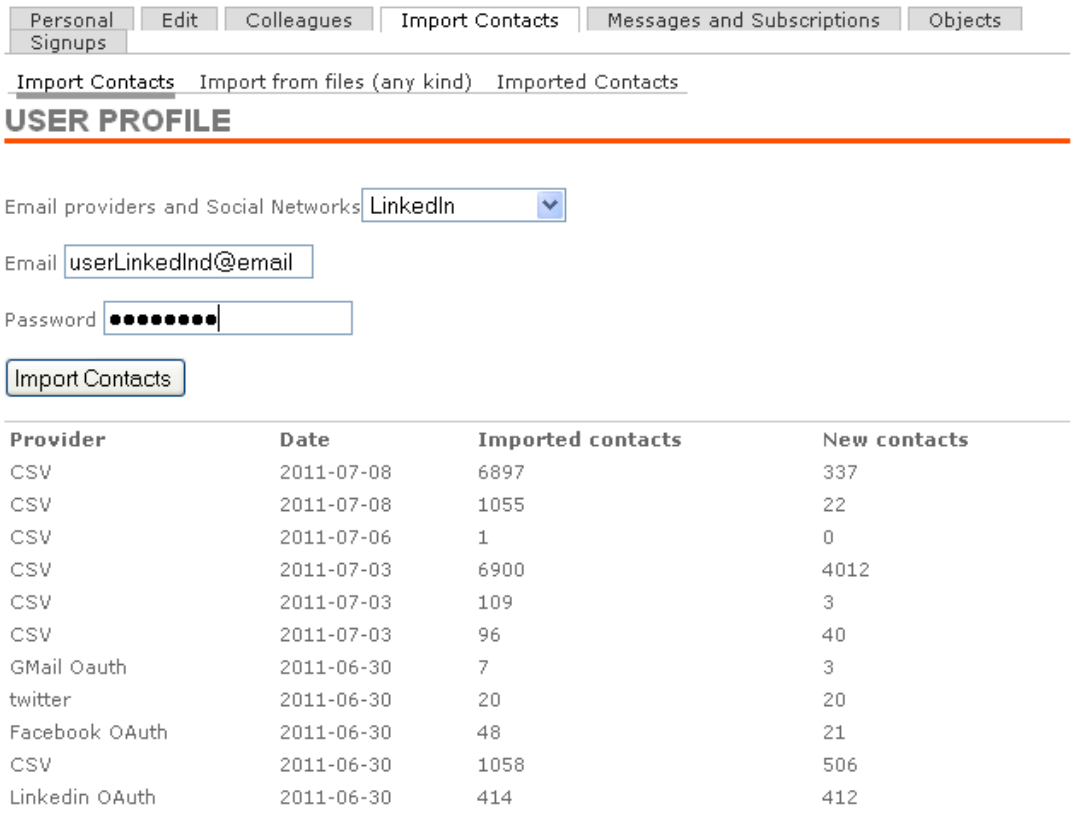

Figure 6.6: Import of contacts from e-mail (APRE Toscana interface).

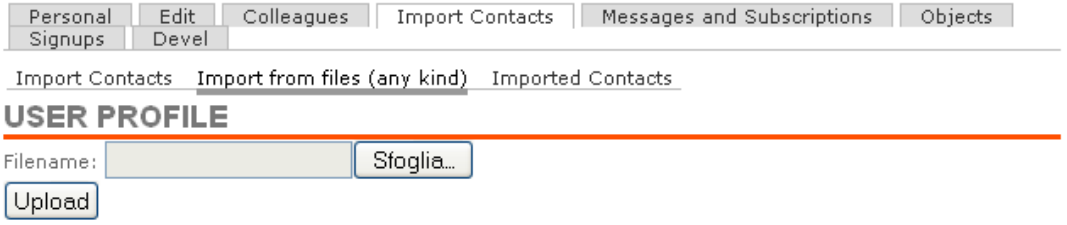

Figure 6.7: Import of contacts from file (APRE Toscana interface).

Figure 6.8: Visualisation of the imported contacts APRE Toscana interface).

6.6.1.Dependencies

Internal component	Type (drupal module, AXCP rule, Java Servlet, Other)	Notes
OpenInviter	Custom module	

Third-party software	Notes
OpenInviter	Php library

Conclusions

In this thesis, the treatment of knowledge has been analysed starting from a theoretical point of view, passing from the ICTs analysis and finally designing and implementing a framework capable of managing the flow of knowledge in all its complexity and entirety. In order to do this, the knowledge concept, its features, the models for the knowledge treatment have been analysed. It has been understood how the knowledge can be treated, investigated and then formalized and shared, moreover it has been clarified that another important aspect that has to be taken into account for the knowledge diffusion is related to the methods and models of interaction and communication among people. Most of the studies realised in the course of history describe and formalise the knowledge treatment as a set of ideas that born, are spread and grow up among through people interactions.

On the other side, also the Information and Communication Technologies that can be applied to in the management of the knowledge flow have been analysed: ITCs on how the information can be acquired, retrieved, shared, managed, discovered, use to extract new knowledge, protected, etc. have been identified, analysed, described in details and finally compared among them.

This technical comparison has been useful to understand that the best way to cover all the aspects related to the knowledge management is through a Best Practice Network (BPN): its features are such as to provide to its users services centered both on the content management and on social interactions. Once detected the BPN as the best way to achieve an environment that integrates a range of tools and services for the development of knowledge, you pass to the general description of the architecture that was first designed and then implemented. The architecture developed offers to its users a large set of functionalities, so it can be easily customised and applied in many different contexts. To give credit to this statement, three different scenarios in which the BPN has been successfully applied have been described: Mobile Medicine solution, European Collected Library of Artistic Performance (ECLAP, ICT PsP project of the European Commission), Sportello APRE Toscana.

Mobile Medicine Portal offers both an health care content distribution service and a collaborative networking portal, for discussion on procedures and content, thus suggestions are provided at the same time on PC and Mobiles. The possibility to have information on mobiles is relevant to solve and support emergency occurrences and first aid treatments in hospitals.

The ECLAP European Project develops a Best Practice Network acting as a content aggregator, making use of advanced database and delivery tools for the production and dissemination of the rich multilingual European heritage.

The Sportello APRE Toscana promotes Italian and Tuscan participation in the European Union's Programmes of research, development an Regional APRE Committee, supports, cooperates and provides the strategic policy for APRE's development activities.

Bibliography

[Abb] Nicola Abbagnano, Giovanni Fornero. Filosofi e Filosofie nella Storia, vol I. Ed. Paravia.

[AnBrGaJo] Anagnostopoulos A., Broder A.Z., Gabrilovich E., Josifovski V., Riedel L.. Just-in- Time Contextual Advertising. CIKM‘07, November 6–8, 2007, Lisboa, Portugal.

[APRE] APRE- Agenzia per la Promozione della Ricerca Europea. http://www.apre.it/

[APRETos] Sportello APRE Toscana. Available at: http://bpnet.apretoscana.org/

[Axmedis] Axmedis European project, Automated Production of Cross Media Content for Multichannel Distribution. Information available at: http://www.axmedis.org

[BaBlHe] P. Van Baalen, J. Bloemhof-Ruwaard, E. Van Heck. “Knowledge Sharing in an Emerging Network of Practice: The Role of a Knowledge Portal”. European Management Journal Vol. 23, No. 3, pp. 300–314, 2005. doi:10.1016/j.emj.2005.04.008

[BaMuSe] Bartz K., Murthi V., Sebastian S. Logistic Regression and Collaborative Filtering for Sponsored Search Term Recommendation. EC‘06, June 11–15, 2006, Ann Arbor, Michigan.

[BBCFNP] P.Bellini, I.Bruno, D.Cenni, A.Fuzier, P.Nesi, M.Paolucci. Mobile Medicine: semantic computing management for health care applications on desktop and mobile devices. In press on 'Multimedia tool and applications', Springer. http://www.springerlink.com/content/q8512555u0j00584

[BCFNPPa] P. Bellini, D. Cenni, A. Fuzier, P. Nesi, G. Pantaleo, M. Paolucci. ‘Metrics for best practice networks analysis’. The 16th International Conference on Distributed Multimedia Systems, Hyatt Lodge at McDonald's Campus, Oak Brook, Illinois, USA. October 2010.

[BeBFNP] P.Bellini, I.Bruno, A.Fuzier, P.Nesi, M.Paolucci, 'Semantic processing and management for Multichannel Cross Media Social Networking'. NEM Conference Saint Malo, France, Septemper 2009.

[BeBN] Bellini P., Bruno I. and Nesi P. A Distributed Environment for Automatic Multimedia Content Production based on GRID, in Proc. AXMEDIS 2005, Italy, 30 Nov. - 2 Dec., pp.134-142, IEEE Press.

[BeBN06] Bellini P., Bruno I., Nesi P. A language and architecture for automating multimedia content production on grid, Proc. of the IEEE Int. Conference on Multimedia & Expo (ICME 2006), IEEE Press, Canada, 9-12 July, 2006.

[BeBNR] P. Bellini, I. Bruno, P. Nesi, D. Rogai. MPEG-21 for the multi-format and multi-channel content production, protection and distribution, Proc. of the Int. Broadcasting Conf., IBC 2007, Amsterdam, Sept. 2007.

[BeBNS] Bellini P., Bruno I., Nesi P., Spighi M. Intelligent Content Model Based on MPEG-21, AXMEDIS '08. International Conference on IEEE press Automated solutions for Cross Media Content and Multi-channel Distribution, 17-19 Nov. 2008, pp:41-48.

[BelGri] Petros Belsis, Stefanos Gritzalis. Distributed autonomous Knowledge Acquisition and Dissemination ontology based framework. Workshop on Ontology and Enterprise Modelling: Ingredients for Interoperability. Proceedings . In Conjunction with 5th International Conference on Practical Aspects of Knowledge Management, December 2, 2004. Vienna, Austria. Edited by Harald Kühn BOC Information Systems GmbH.

[BeNePaSe] P.Bellini, P.Nesi, M.Paolucci, M.Serena. ”Models and tools for aggregating and annotating content on ECLAP” (short paper). ISM2011, IEEE International Symposium on Multimedia.

[BeNePaSeDMS] P.Bellini, P.Nesi, M.Paolucci, M.Serena. MyStoryPlayer, Collections, Playlists: Content aggregator tools in Eclap Project. Demo paper at DMS 2011, Firenze 18-20 Agosto, 2011.

[BenFensPer] V. Richard Benjamins, Dieter Fensel, Asunción Gómez Pérez. “Knowledge Management through Ontologies”.Proceedings of the Second International Conference on Practical Aspects of knowledge Management, 1998. Available at: http://citeseer.ist.psu.edu/viewdoc/summary?doi=10.1.1.42.2746

[BeNR] Bellini P., Nesi P., Rogai D. Exploiting MPEG-21 File Format for Cross Media Content. 13th Int. Conf. on Distributed Multimedia Systems (DMS), San Francisco Bay, USA, 6-8 September 2007.

[BreGal] Brent Gallupe. “Knowledgemanagement systems: surveying the landscape”. International Journal of Management Reviews, Volume 3, Issue 1, pages 61–77, March 2001. Available at: http://onlinelibrary.wiley.com/doi/10.1111/1468-2370.00054/abstract

[BrFoJoRi] Broder A., Fontoura M., Josifovski V., Riedel L.. A Semantic Approach to Contextual Advertising. Proc. of 30th annual int. ACM SIGIR conf. on Research and development in information retrieval, SIGIR ‘07.

[BusDic] Business Dictionary. Available at: http://www.businessdictionary.com/definition/knowledge.html

[CiMuPl] Ciaramita M., Murdock V., Plachouras V.. Semantic Associations for Contextual Advertising. Journal of Electronic Commerce Research, VOL 9, NO 1, 2008.

[Cin] Maria Cinque. Agire Creativo – Teoria, formazione e prassi dellínnovazione personale. Fondazione Rui.

[Cordis] CORDIS: http://cordis.europa.eu/home_en.html

[CuVZee] S. Cummings, A. Van Zee. “Communities of practice and networks: reviewing two perspectives on social learning”. Cummings, S. and A. van Zee, 2005. Communities of practice and networks: reviewing two perspectives on social learning KM4D Journal 1(1): 8-22.

[DC] Dublin Core Metadata Initiative. Available at: http://dublincore.org/

[DruInv] Drupal Invite. Avalable at: http://drupal.org/project/invite.

[DruMime] Drupal Mime Mail. Available at: http://drupal.org/project/mimemail

[DruNot] Drupal Notifications. Available at: http://drupal.org/project/notifications.

[DruOg] Drupal Organic Groups. Available at: http://drupal.org/project/og

[DruOgNot] Drupal Organic Groups Notifications. Available at: http://drupal.org/project/og

[Drupal] Drupal, Open Source CMS. Site: http://drupal.org.

[DruTax] Drupal Taxonomy Context. Available at: http://drupal.org/project/taxonomy_context

[DruTaxImp] Drupal Taxonomy Import/Export via xml. http://drupal.org/project/taxonomy_xml

[EC-DE3.1] European Collected Library of Artistic Performance- DE3.1 Infrastructure: ingestion and processing content and metadata. Available at: http://www.eclap.eu/drupal/?q=en-US/home&axoid=urn:axmedis:00000:obj:a345a84f-6fdf-4f84-a412-

88094ce363e2&section=search_base_result&cd=26&qid=4ca3bbb8b684f61187313604bbf937a5&n=93

[EC-DE3.3.1] EUROPEAN COLLECTED LIBRARY OF ARTISTIC PERFORMANCE - DE3.3.1. Infrastructure: Content and Metadata processing and semantification.

[EC-DE6.2.2] ECLAP Specific Services. DE 6.2.1.

[Eclap] European Collected Library of Artistic Performance. Platform available at: http://www.eclap.eu/

[EclapParts] Eclap Partners. List and information on Eclap Partners available at: http://www.eclap.eu/drupal/?q=node/3578

[ECToU] Eclap Terms of use. Available at: http://bpnet.eclap.eu/drupal/?q=en-US/node/3730

[Esa] Esanum, communio adiuvat. Available at: http://www.esanum.it

[eu_agreement] 'Europeana Data Provider & Aggregator Agreements License for use of content with partners and aggregators of content for Europeana Content Use Contract', http://group.europeana.eu/c/document_library/get_file?uuid=1c287538-d3c9-4843-9992-c4539f16aec0&groupId=10602. Available also on: http://bpnet.eclap.eu/drupal/?q=en-US/home&axoid=urn:axmedis:00000:obj:067aa275-cb89-4c10-8b09-eb50a635246e&section=search_base.

[eu_rights] 'Guidelines for the europeana:rights metadata element', v4.0 - 20/11/2010, http://version1.europeana.eu/c/document_library/get_file?uuid=06e63d96-0358-4be8-9422-d63df3218510&groupId=10602

[Europeana] European Digital Library. Site available at: http://www.europeana.org

[EvLaLe] Everitt, B., Landau, S., Leese, M. Cluster analysis, 4th edition, London, Arnold, 2001.

[Facebook] Facebook: http://www.facebook.com

[FioNe] Fioravanti F., Nesi P., "Estimation and Prediction Metrics for Adaptive Maintenance Effort of Object-Oriented Systems", IEEE Transactions on Software Engineering, IEEE Press, USA, 2001.

[FirMcE] Joseph M. Firestone, Mark W. McElroy. "Doing knowledge management", Learning Organization, (2005). The, Vol. 12 Iss: 2, pp.189 – 212. Available at: http://www.emeraldinsight.com/journals.htm?articleid=1464395&show=abstract

[Flickr] Flickr: http://www.flickr.com

[FRBR] Functional Requirements for Bibliographic Records - FRBR.

[Friendster] Friendster: http://www.friendster.com

[HatGalHua] Afsoun Hatami, Robert D. Galliers, Jimmy Huang. "Exploring the Impacts of Knowledge (Re)Use and Organizational Memory on the Effectiveness of Strategic Decisions: A Longitudinal Case Study". Proceedings of the 36th Hawaii International Conference on System Sciences (HICSS'03) 0-7695-1874-5/03 $17.00 © 2002 IEEE. Available at: http://csdl.computer.org/comp/proceedings/hicss/2003/1874/04/187440110b.pdf

[HogEdv] C. Högberg, L. Edvinsson. "A design for futurizing knowledge networking". Journal of Knowledge Management Volume 2, Number 2 December 1998. Avilable at: http://www.emeraldinsight.com/journals.htm?articleid=883664&show=abstract

[Holl] David Hollingsworth. Workflow Management Coalition The Workflow Reference Model. Document NumberTC00-1003 , Document Status - Issue 1.1, 19-Jan-95.

[HRTS] Hart J., Ridley C., Taher F., C. Sas, A. Dix. Exploring the Facebook Experience: A New Approach to Usability. NordiCHI 2008: Using Bridges, 18-22 October, Lund, Sweden.

[HTML] HyperText Markup Language – HTML.

[Hunt] Darwin P. Hunt. "The concept of knowledge and how to measure it. Journal of Intellectual Capital". Volume: 4 Issue: 1 2003. Available at: http://andrewvs.blogs.com/usu/files/p100.pdf

[ICD] International Classification of Diseases (ICD). Available at: http://www.who.int/classifications/icd

[KraNac] G. Kruitbosch, F. Nack. Broadcast Yourself on YouTube – Really? HCC'08, October 31, 2008, Vancouver, Canada.

[KrNa] Kruitbosch, G., Nack, F. (2008) Broadcast Yourself on YouTube – Really? HCC'08, October 31, 2008, Vancouver, British Columbia, Canada.

[LastFm] LastFm: http://www.lastfm.it

[LHJYPR] Lee, J., Hwang, S.O., Jeong, S.-W., Yoon, K.S., Park, C.S., Ryou, J.-C., "A DRM Framework for Distributing Digital Contents Through the Internet", ETRI Journal, Vol.25, N.6, pp.423-435, December 2003.

[LHJYPR] Lee, J., Hwang, S.O., Jeong, S.-W., Yoon, K.S., Park, C.S., Ryou, J.-C.,. A DRM Framework for Distributing Digital Contents Through the Internet. ETRI Journal, Vol.25, N.6, pp.423-435, December 2003.

[LinkedIn] LinkedIn. http://www.linkedin.com

[Lucene] Apache, Lucene. Site: http://lucene.apache.org/java/docs/index.html

[MARC] MAchine-Readable Cataloging. Site: http://www.loc.gov/marc/

[MBra] Malin Brännback. "R&D collaboration: role of Ba in knowledgecreating Networks". Knowledge Management Research & Practice (2003) 1, 28–38 & 2003.

[METS] Metadata Encoding & transmission Standard - METS. Site: http://www.loc.gov/standards/mets/

[MeySug] Bertolt Meyer, Kozo Sugiyama. "The concept of knowledge in KM: a dimensional model". Journal of Knowledge Management. Volume: 11 Issue: 1 2007. Available at: http://www.emeraldinsight.com/journals.htm?issn=1367-3270&volume=11&issue=1&articleid=1595395&show=html

[MIME] Multipurpose Internet Mail Extensions – MIME

[MiRuSlWi] Michalowski W., Rubin S., Slowinski R., Wilk S., Mobile clinical support system for pediatric emergencies Source, Decision Support Systems, Elsevier Science Publishers B. V. Amsterdam, The Netherlands, Volume 36 , Issue 2 (October 2003), Pages: 161 – 176.

[Mit-Ser-Ce] WP6.4 Slides - ECLAP Internal meeting, Rome 19 October 2011. http://bpnet.eclap.eu/drupal/?q=en-US/home&axoid=urn:axmedis:00000:obj:0e750d66-f7a1-4c11-82a2-c8192bcc026a&section=search_base&cd=1&qid=2e250190d7d8942580aefdeb0988ed65&n=13

[MobMed] University of Florence and Careggi Hospital. Mobile Medicine Project. Available at: http://mobmed.axmedis.org/

[MobMedApp] DISIT, DSI Department of University of Florence. Mobile Medicine iPhone App. Available at: http://itunes.apple.com/en/app/mobilemedicine/id359865882

[MobMedHelp] DISIT, DSI Department of University of Florence. Mobile Medicine Help. Available at: http://mobmed.axmedis.org/drupal/?q=en-US/help

[MobMedUM] Mobile Medicine User Manual. Available at: http://mobmed.axmedis.org/Mobile-Medicine-Manual.pdf

[MOC] MyOpenCare – Share Healt Contents. Available at: http://www.myopencare.com

[MODS] Metadata Object Descritpion Schema. Site: http://www.loc.gov/standards/mods/

[Moodle] Moodle Course Management System. Site: http://moodle.org

[MPEG-7] MPEG-7, "Multimedia Content Description Interface".

[MySpace] MySpace: http://www.myspace.com

[NaRaCh] A. Nazir, S. Raza, C. Chuah. Unveiling Facebook: A Measurement Study of Social Network Based Applications. IMC'08, October 20–22, 2008, Vouliagmeni, Greece.

[NeRoV] Nesi, P.; Rogai, D.; Vallotti, A. A Protection Processor for MPEG-21 Players. In Multimedia and Expo, 2006 IEEE International Conference. Page 1357-1360. 9-12, July 2006, Toronto, Canada.

[NLMMesh] MeSH Tree Structures, U.S. National Library of Medicine National Institutes of Health. Available at: http://www.nlm.nih.gov/mesh/2009/mesh_trees/trees.html

[NonKatDaS] Ikujiro Nonaka, Katsuhiro Umemoto and Dai Senoo. "From Information processing to knowledge Creation: A Parasdigm Shift in Business Management". Technology In Society, Vol. 18. No. 2, pp. 203-218, 1996.

[NonTak] Nonaka Ikujiro, Takeuchi Hirotaka. "The Knowledge-Creating Company". Oxford University Press, Oxford, New York, 1995.

[NonToyKon] Ikujiro Nonaka, Ryoko Toyama and Noboru Konno. "SECI, Ba and Leadership: a Unified Model of Dynamic Knowledge Creation". Pergamon Long Range Planning 33 (2000) 5±34.

[OAI-ORE] Open Archives Initiative Object Reuse and Exchange, OAI-ORE. Site: http://www.openarchives.org/ore/

[Orkut] Orkut: http://www.orkut.com

[OWL] Web Ontology Language - OWL.

[OxfDic] Oxford Dictionaries. Available at: http://oxforddictionaries.com/

[Pao] Michela Paolucci. Una soluzione Innovativa per la modellazione e gestione distribuita di documenti giuridici in rete. Bachelor Thesis, University of Florence, Faculty of engineering, Department of electronics and telecommunications, AA2005/2006.

[PiTa] Pirr, G.; Talia, D., An approach to Ontology Mapping based on the Lucene search engine library, 18th International Conference on Database and Expert Systems Applications, 2007. DEXA, Volume , Issue , 3-7 Sept. 2007 Page(s):407 – 411.

[PM] PubMed - US National Library of Medicine National Institutes of Health. Available at: http://www.ncbi.nlm.nih.gov/pubmed

[PrGrHe] Protogerakis, M.; Gramatke, A.; Henning, K., A System Architecture for a Telematic Support System in Emergency Medical Services, ICBBE 2009. 3rd International Conference on Bioinformatics and Biomedical Engineering , 2009, Issue , 11-13 June 2009 Page(s):1 – 4.

[QuiLeJon] Paul Quintas, Paul Lefrere, Geoff Jones. "Knowledge management: A strategic agenda". Long Range Planning, Volume 30, Issue 3, June 1997, Pages 385-391.

doi:10.1016/S0024-6301(97)90252-1. Available at: http://www.sciencedirect.com/science/article/pii/S0024630197902521

[RDF] Resource Description Framework - RDF.

[RicRic] John L. Rice & Bridget S. Rice. "The Applicability of the SECI Model to Multi-Organizational Endeavours: an integrative review".

[RoLe] Rousseeuw P. J., Leroy A. M., Robust regression and outlier detection, John Wiley & Sons, Inc., New York, NY, 1987.

[RonYou] Ron Young. "Knowledge Management - Back to Basic Principles". Article on Knowledge-management-online.com. Available at: http://www.knowledge-management-online.com/knowledge-management-back-to-basic-principles.html

[Ruff] M.Ruffolo. Tecnologie informatiche per la gestione della conoscenza e del capitale intellettuale. Riferimenti metodologici e studi di caso per la gestione e la valorizzazione. http://dns2.icar.cnr.it/ruffolo/files/Dispensa%20-%20Ruffolo.pdf

[SBN] OPAC SBN - Catalogo del Servizio Bibliotecario Nazionale.

[SCORM] Sharable Content Object Reference Model - SCORM

[SeVoKBa] A. Seufert, G.Von Krogh A. Bach. "Towards knowledge networking". Journal of Knowledge Management, Volume 3 . Number 3 . 1999 . pp. 180-190. MCB University Press . ISSN 1367-3270.

[SGML] Standard Generalized Markup Language - SGML.

[SMIL] Synchronized Multimedia Integration Language - SMIL. http://www.w3.org/TR/REC-smil

[Sparql] Simple Protocol and RDF Query Language – SPARQL.

[SulHaw] Suliman Al-Hawamdeh. "Knowledge management: re-thinking information management and facing the challenge of managing tacit knowledge". Information Research, 8(1), paper no. 143, (2002) . Available at: http://InformationR.net/ir/8-1/paper143.html

[TimKuz] Timo Kucza. "Knowledge Management Process Model". VTT Publications n. 455, TECHNICAL RESEARCH CENTRE OF FINLAND, ESPOO 2001.

[UML] Unified Modeling Language - UML.

[XML] eXtensible Markup Language – XML.

[XuWu] Xui R., Wunsch D. C. II. Clustering, John Wiley and sons, USA, 2009.

[YiGoCa] Yih W., Goodman J., Carvalho V.R.. Finding Advertising Keywords on Web Pages. WWW 2006, May 23–26, 2006, Edinburgh, Scotland.

[YouTube] YouTube: http://www.youtube.com

[ZhSuPlNa] Zhang Y., Surendran A.C., Platt J.C., Narasimhan M.. Learning from Multitopic Web Documents for Contextual Advertisement. KDD'08, August 24–27, 2008, Las Vegas, Nevada, USA.

Appendix I: Other Publications

- P.Bellini, I.Bruno, D.Cenni, P.Nesi, M.Paolucci, M.Serena. Semantic Model for Cultural Heritage Social Network and Cross Media Content for Multiple Devices. AI*IA, workshop beni culturali, Palermo 2011.

- P.Bellini, P.Nesi, M.Paolucci, M.Serena. “ECLAP: European Collected Library of Performing Art”. Demo paper at NEM 2011, Torino 27-29 settembre 2011

- P.Bellini, I.Bruno, D.Cenni, P.Nesi, M.Paolucci, 'Personal Content Management on PDA for healt Care Applcations', 2009 IEEE international workshop on Semantic Computing and Mulimedia Systems, Berkeley CA, USA. 14-16 September 2009.

- F. Frosini, N. Mitolo, P. Nesi, M. Paolucci, "Collaborative Solution for Music Education", Automated solutions for Cross Media Content and Multi-channel Distribution, 2008 AXMEDIS '08, International Conference on IEEE press, 17-19 Nov. 2008 Page(s):71 – 78.

- P. Bellini, F. Frosini, N. Mitolo, P. Nesi, M. Paolucci, “Collaborative Working for Music Education”, 8th International conference on new Interfaces for musical expression, 4th IMAESTRO workshop on Technology Enhanced Music Education, 5-6 June, 2008, Genova, Italy.

- P. Bellini, F. Frosini, N. Mitolo, P. Nesi, M. Paolucci, “Music Representation with MPEG-SMR”, 8th International conference on new Interfaces for musical expression, 4th IMAESTRO workshop on Technology Enhanced Music Education, 5-6 June, 2008, Genova, Italy.

- Prof. Franco Pirri, Ing. Michela Paolucci, Ing. Davide Chini, Ing. MariaChiara Pettenati, Ing. Samuele Innocenti, "InterDataNet: Interoperability Middleware Infrastructure to Support Collaborative Creation and Management of Ofiicial Documents in e-Government Processes", 41th Hawaii International Conference on System Sciences.

- Ing. Michela Paolucci, Ing. Davide Chini, Ing. Samuele Innocenti, Ing. MariaChiara Pettenati, "Use of an Information and a Middleware infrastructure for Collaborative creation and management of official documents in e-Government back-office procedures", Egov 2007 – the International Conference of the EGOV-Society”.

Appendix II: ECLAP Questionnaire

To realize an appropriate model for the Content and Metadata Ingestion on the ECLAP Portal, a questionnaire has been distributed to all Content Partners regarding the following main topics: General contacts; Content and metadata Upload methods; Metadata standards and formats; IPR on content (licenses, permissions, etc.); Collection topics; etc. In details, (see also [EC-DE3.1]):

- **Contacts**: CP role (university, etc.), person responsible for ingestion, person responsible for IPR, etc.
- **Content type**: video, audio, text, image, html, animation, etc.
- **Content format**: it depends from the type (e.g. for the videos: flv, mov, avi, MPEG1, MPEG2, SWF, RM, WMV, etc.)
- **Content Upload method**: Hard Disk (the CP can deliver contents stored into a Hard Disk), http (the CP will provide a list of HTTP URL for each content available on a own or hosted Web server), ftp (the CP will provide a list of FTP URL and related access credential for each folder containing contents available on a own or hosted FTP server), OAI-PMH (the CP can provide OAI-PMH metadata ingestion containing valid URLs for retrieving content), etc.
- **Content IPR status specification**: Public Domain, IPR not cleared, content associated to a specific license (that has to be provided by the CP), Orphan Work, etc.
- **IPR standard / registration**: if this information is registered, a standard can be used (for instance: MPEG-21 REL, OMA DRM), or it can be collected in various fields in CP's metadata schema.
- **Content Rights holders agreements**: fundamental to establish if CP's institution got an agreement with a collecting societies of rights holders regarding the management of royalties on the content that it can provide to ECLAP (chooses between: 'No, but we DO expect to have to pay royalties for ECLAP content', 'No, and we DONT expect to have to pay royalties for ECLAP content', 'Yes, and we DON'T expect to have to pay royalties for ECLAP content', 'Yes, and we DO expect to have to pay royalties for ECLAP content')
- **Content Technological protection**: a CP has to declare the technology eventually used to protect its digital content for online use (digital watermarking, scarring, low resolution, etc).
- **Content Restrictions/Permissions**: they depend on the content. Each CP has to communicate to which ECLAP users and under what conditions the content will be accessible/downloadable, etc.
- Metadata upload method (same as what described for content upload method): Hard disk, http, ftp, OAI-PMH, etc.

- Metadata standard: if present the CP can select among CDWA, CIDOC-CRM, Dublin Core, EAD, FRBR, ISAD(G), MAB, MARC, METS, MIDAS, MODS, museumdat, Object ID, SPECTRUM, TEI, VRA, etc.
- Metadata format: the CP has to specify the format in which the metadata are written (sending the metadata files, possibly in xml if the exist).
- Metadata Language
- General description
- Topics
- Time period

Specific questions have been submitted to the Content Partners to specify the type of permissions they would like to apply to their content. These information are fundamental and reflect the permissions for the IPR Wizard. Access permissions to be applied depend on users roles and on Content types. Each permission is associated to a question on the questionnaire:

PDF/etc.

Permission applied	Question for each Content Set	Possible answers
Document access (PC and Mobile)	'Which are the users that can access (play on web page and download) the content of this Content Set?‘	• All ECLAP Users • Only users registered to a content group • None

AUDIO

Permission applied	Question for each Content Set	Possible answers
Audio download-PC	'Which are the users that can download the content of this Content Set in Low Definition (LD) and Medium Definition (MD) on PC?'	• All ECLAP Users • Only users registered to a content group • None
Audio play-PC	'Which are the users that can play on line (streaming or progressive) the content of this Content Set in Low Definition (LD) and Medium Definition (MD) on PC?'	• All ECLAP Users • Only users registered to a content group • None
Audio download-mobile-Browser	'Which are the users that can download the content of this Content Set from mobile via Browser?'	• All ECLAP Users • Only users registered to a content group • None
Audio play-mobile-Browser	'Which are the users that can play (streaming or progressive) the content of this Content Set on	• All ECLAP Users • Only users registered to a

	mobile via Browser?'	content group • None
Audio download-mobile-Apps Content Organizer	'Which are the users that can download the content of this Content Set from mobile via ECLAP applications?'	• All ECLAP Users • Only users registered to a content group • None
Audio play-mobile-Apps Content Organizer	'Which are the users that can play (streaming or progressive) the content of this Content Set on mobile via ECLAP Applications?'	• All ECLAP Users • Only users registered to a content group • None

VIDEO

Permission applied	Question for each Content Set	Possible answers
Video download PC HD	'Which are the users that can play the content of this Content Set in High Definition Quality (HD), on PC?'	• All ECLAP Users • Only users registered to a content group • None
Video play PC HD	'Which are the users that can play the content of this Content Set in High Definition Quality, on PC?'	• All ECLAP Users • Only users registered to a content group • None
Video download-PC- LD and MD	'Which are the users that can download the content of this Content Set in Low Definition (LD) and Medium Definition (MD) on PC?'	• All ECLAP Users • Only users registered to a content group • None
Video play-PC- LD and MD	'Which are the users that can play online (streaming or progressive) the content of this Content Set in Low Definition (LD) and Medium Definition (MD) on PC?'	• All ECLAP Users • Only users registered to a content group • None
Audio download-mobile-Browser	'Which are the users that can download the content of this Content Set from mobile via Browser?'	• All ECLAP Users • Only users registered to a content group • None
Audio play-mobile-Browser	'Which are the users that can play (streaming or progressive) the content of this Content Set on mobile via Browser?'	• All ECLAP Users • Only users registered to a content group • None
Audio download-mobile-Apps Content Organizer	'Which are the users that can download the content of this Content Set from mobile via	• All ECLAP Users • Only users registered to a

	ECLAP applications?'	content group • None
Audio play-mobile-Apps Content Organizer	'Which are the users that can play (streaming or progressive) the content of this Content Set on mobile via ECLAP Applications?'	• All ECLAP Users • Only users registered to a content group • None