INTRODUCTION

This research aims at assessing argumentation as a means of mediating large-scale, online deliberation on a wicked problem. Collective online deliberation entails the discussion and exploration of possible solutions on a question in hand. It has been rising as a way of leveraging the collective intelligence of communities while gathering geographically dispersed people from where they are. Although various existing methods such as wikis and forums can be used for an online deliberation, their use has been mostly limited to the questions easy to handle, that is, tame problems.

According to Rittel & Weber (1973), solving a problem of mathematics such as an equation or a problem of chemistry such as obtaining a compound through chemical reactions, provide clear missions to accomplish, letting one to assess whether a solution considered solves the problem. Professionalism in diverse fields of science and engineering has been a way of dealing with tame problems of this nature. However the "rational" methods that scientists and engineers use for solving tame problems may not work for many issues of modern times such as societal problems. Problem solving may actually entail a planning process composed of a series of actions yielding an output which may initiate another process of problem solving. Therefore these "wicked" problems require an interconnected set of actions making it difficult for one to locate the problem within the network and to decide from where to start to solve the problem. This phenomenon requires internalizing what was once external to us. Characteristics of wicked problems can be summarized as follows:

- Wicked problems cannot be formulated easily. The search for the necessary information to solve the problem requires an idea on possible paths of solution. For instance global warm'ng

problem can be related to various topics ranging from pollution to the lack of regulation to the absence of authorities capable of setting an efficient environmental policy. Defining the problem is therefore, somehow, the same as finding the solution.

- The aforementioned interconnected topics constitute a network of problems. In the absence of limitless resources such as time, financing and expertise exploring the whole network is infeasible.
- Judgments on wicked problems cannot be as true or false as for tame problems. Solutions to wicked problems are rather good or bad.
- While solutions to tame problems can be tested under controlled conditions, this does not apply for wicked problems. It is practically impossible to observe the consequences through the whole chain of actions. By the same token, consequences of actions are irreversible for wicked problems. Once they are done it is not possible to go back through the chain of actions and observe the process whereas the solutions to tame problems can be repeated. This also implies that a wicked problem solver cannot fail. Because in case of failure, setting up the system back to its previous state is impossible.
- Wicked problems are always different from each other requiring customized solutions for each problem.
- A wicked problem can be a successor/predecessor of another. For instance global warming problem can be seen as a result of the pollution problem while pollution can be related to the industrialization. Therefore a problem can be explained in various ways.

Requiring the exploration of a solution space that may easily span various fields while being subject to constraints such as time and expertise, it is not uncommon that several people are involved in solving a wicked problem. At the time of Internet, many computer supported tools

provide a convenient way to gather people without moving them from their places. These tools can be grouped into two as:

1. *Group Decision Support Systems*: Examples include decision support systems and online Delphi method applications. They favor a high level of structure for better organization. As a tradeoff they cannot be used at large scales. Actually they were only tried in laboratory conditions with a trainer acting as a facilitator.
2. *Traditional Online Tools*: They are tools such as forums, wikis, instant messaging etc. There is evidence that they can be used for collaborative purposes at large scales as success stories such as Wikipedia show. The loose structure of tools let prospective users to adopt them considerably easily spreading their use. In spite of having a proven record in supporting knowledge sharing, they accumulate knowledge instead of adding pieces of knowledge in a meaningful way. They do not let evaluate a certain piece of knowledge and find relevant information in the mass. Thus the achievements obtained in online knowledge sharing do not necessarily mean success at the deliberation of wicked problems. Many examples of these tools do not qualify as good media of deliberation because of the problems such as:

- Redundancy: Many tools suffer from redundant information hampering the conciseness, e.g. forums
- The lack of navigability: While the tools may let covering the whole related content, they may have problems in showing the connection between related pieces of information e.g. wikis.
- Conflicts: Opposite views may produce clashes that block conversations e.g. some controversial Wikipedia articles are edited very frequently by people representing different

views. Therefore instead of developing through different ideas, the article goes back and forth between versions representing a single view.

Thus there is need for better ways to let large groups deliberate on a wicked problem through online methods. An alternative tested here is argumentation. It is based on the idea that every discourse is basically composed of a claim and premises that constitute the ground for that claim. Argumentation supports critical thinking by implicitly persuading people to give evidence and logic in their contributions. The arguments are easily represented by an argument map that provides a concise way of representing contributions facilitating finding the related piece of information while decreasing redundant entries.

I.1) RESEARCH QUESTION

Present study explores argumentation-based tools as a way of conducting large scale online deliberations. The assumption is that in case argumentation is capable of surmounting the obstacles on redundancy, navigability and conflicts, it should be good at enabling communication across time and people. However, argumentation suffers from the typical problems of group decision support system tools described above and has not received widespread diffusion as an online tool. This has raised concerns on the capability of argumentation to act as an effective mediator, notwithstanding the remarkable advantages that are expected from its use. Thus in this work the aim is to investigate to what extent and under which conditions argumentation tools can be used for large-scale deliberation and how their mediating capability can be improved by proper design.

Previous research on the development of online argumentation has privileged the issue of knowledge representation. Therefore the focus has been on the construction of the appropriate

knowledge formats for capturing and displaying user contributions as well-formed argumentation, rather than the effectiveness of mediation in the interaction process leading to the knowledge production in the right format. The assumption is that in order to be qualified as a feasible way of deliberation, argumentation tools should enable mutual understanding through the accumulation of knowledge through time on the top of previous contributions.

Thus, this research aims at assessing the mediation capability of an online argumentation tool in terms of common understanding construction through argument-based conversations. With these in mind, key questions guiding this research are:

1. How can the development of common understanding be materialized and measured for argumentation?
2. How does common understanding on different discussion topics develop among deliberators?
3. How does the presence of common understanding can be helpful in designing future tools?

I.2) LITERATURE OVERVIEW

Providing an answer to the questions above implies constituting a theoretical basis for the assessment of other deliberation media, a benchmark for new tool designers focusing on solving wicked problems. In the process of solving such problems a group of people can come up with better results than isolated individuals would, whereas some difficulties may arise as shown by the mixed results of scholars. Presence of diverse skills and competences may allow a richer problem analysis (S. E. Page, 2008) and increasing productivity (Pelled, Eisenhardt, & Xin, 1999; Vennix, 1996). However the very same diversity may hamper collaboration as posited by the Absorptive Capacity concept (Cohen & Levinthal, 1990). Actually not all collaborations have

been successful (Barron, 2003) as groups face three major obstacles against effective collaboration according to (Sunstein, 2006):

1. Due to social pressure, individuals may hesitate to disclose information.
2. Earlier contributions may have an effect disproportional to their validity in the rest of the deliberation (cascade effects).
3. Group polarization may arise particularly on controversial issues.

A tool designed for an online deliberation is useful to the extent it allows deliberating participants to delve deeper into the discussions, bearing in mind the above obstacles. A long time objective of Information Sciences (IS) research has been gathering geographically dispersed people. Currently available tools are regrouped under three major headings by De Moor & Aakhus (2006):

1. Funneling technologies such as prediction markets and e-voting tools are good at aggregating individual opinions and filtering the most supported view. However they do not provide any means to explore the advantages and disadvantages of available options.
2. Sharing tools (Josang, Ismail, & Boyd, 2007) consist of wikis, blogs, forums etc. These tools have been widely popular. However they suffer from redundant and unsystematic contributions. They are also notorious for disputes such as forum flame wars and wiki edit wars that happen particularly at the discussion of controversial topics.
3. Argumentation is used to a limited extent in comparison with the previous two groups of tools. It is founded on the idea of reaching conclusions through logical reasoning. However when people are involved in a debate, they overvalue the arguments that support their view (confirmation bias). Interestingly, while this flaw may produce individual arguments of poor

quality, it implies an advantage for collective argumentation because people are good at falsifying ideas when they disagree with them (Sperber & Mercier, 2009). This also implies providing arguments of high quality to avoid counter statements. Thus argument has the potential of drawing a productive deliberation out of a controversial topic whereas sharing tools are vulnerable to conflicts.

Toulmin (2003) saw argument as a chain of claims supported by premises. Walton (2006, 1989) suggested a set of schemes where each scheme provided with a question to verify the reliability of the premise. Finally IBIS (Issue-Based Information System, Conklin, 2006) provided a tree like structure organizing argumentation through a set of questions to be answered, options as alternative solutions to these questions and supportive or challenging pro/con arguments on the options. These characteristics allow representing different opinions through an argument map (**Figure 1**).

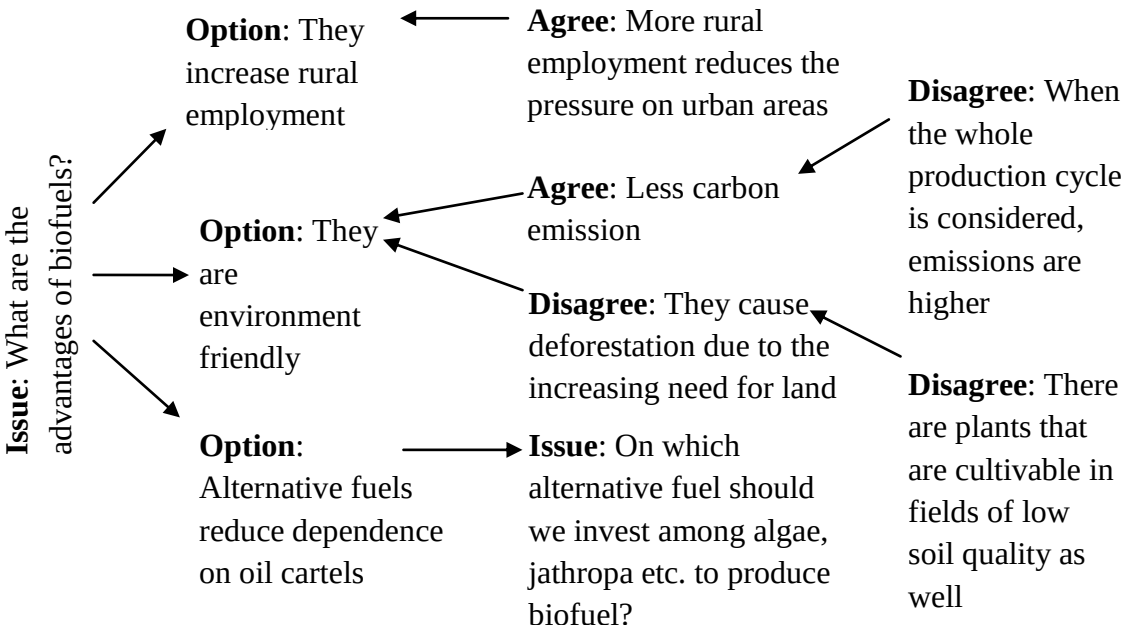

Figure 1 An exemplary argument map

An argument map, the output of deliberations conducted through argumentation has two major functions. First, it is a way of externalizing knowledge as every contribution made is represented through arguments. Second it permits to analyze and explore the once externalized knowledge. According to Clark (1996), the effective construction of common understanding (common ground) is a necessary condition for a successful knowledge transformation in a conversation. However there is not enough evidence that argument-based interaction will bring to the construction of a higher level of common understanding and awareness. With this in mind, the present work studies the deliberation process through a common ground perspective in order to assess if and to which extent interaction mediated by an argumentation tool can make a difference. Common ground stands for mutual knowledge, beliefs and attitudes among participants developed through their interaction into a joint collaborative activity (1996). Referring to the standpoint of conversational analysis, Clark & Schaefer (1989) summarize two conditions for an effective common ground building:

1. *Basis*: There is supposed to be a common ground on which participants agree.
2. *Accumulativeness*: Participants build knowledge on the basis of the presupposed common ground and this depends on making the right addition of knowledge at the right time.

In line with this view, common ground does not only refer to the present status of discussions but also to the process that led to that status. As Deshpande et al. (2005) posit, a shared understanding develops at three steps within a community:

1. At the problem definition stage, people formulate the problem and identify the goals while trying to figure out what others' interests and beliefs are.

2. At a second, individual perspective stage, people exchange their beliefs through three sub-stages, namely elicitation, elimination and externalization. In elicitation, community members try to grasp the context to which their knowledge applies. In elimination, people merge ideas that have the same meaning. Finally in externalization, people establish the relationships among different ideas raised by different community members.

3. The final, team perspective stage, entails the efforts in reflecting on the conclusions derived at stage two and assessing their viabilities. The outcome of this stage may require revising the individual knowledge, creating a continuous loop of common ground building.

In such environment where the common understanding is subject to continuous update, securing the right conveyance of messages is crucial. In order to overcome this difficulty formalisms are suggested (Beers, Boshuizen, Kirschner, & Gijselaers, 2005). Formalisms constrain the way the discussions held in order to structure the way the exchanges are held. The outcome of this is usually the construction of a knowledge object that is a representation of knowledge, facilitating communication while decreasing divergence in discussions. Therefore a knowledge object that is more representative of the deliberation content is also expected to be a better communication mediator.

While the level of complexity of the problem in hand increases, communication risks become more prevalent. This can be offset with a higher level of formalism. The way argumentation categorizes the elements of a discourse is a type of formalism that produces an argument map as a knowledge object. Argumentation is supposed to provide a better knowledge object in comparison with sharing and funneling tools rendering it convenient for building a common

understanding. In other words, it has the potential to combine the advantages of online communication with knowledge objects (**Figure 2**).

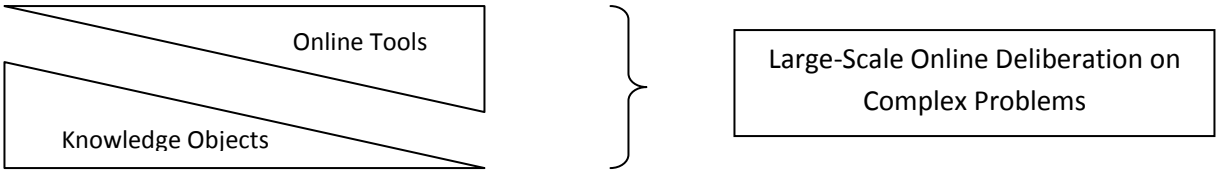

Figure 2

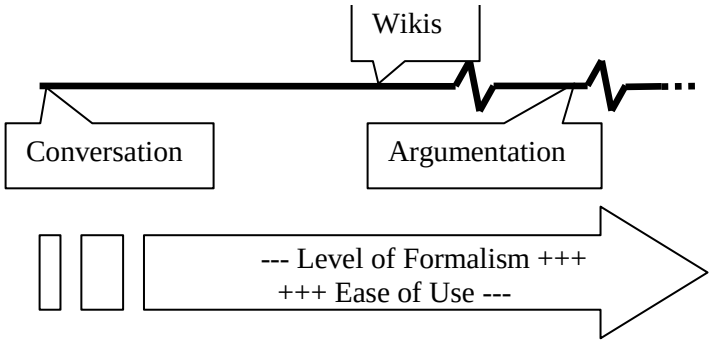

Figure 3 Continuum of Formalism and Ease of Use

Although knowledge object works as a facilitator for the deliberation, a higher level of formalism may impede its advantages in developing common ground (**Figure 3**). First, coercions implied by formalisms steepen the learning curve for the deliberation tool. Second, formal constraints imply a reduction in the communication bandwidth, i.e. some information is lost because the structure implemented does not let it. For instance unlike forums, argumentation does not let the communication of informal content which may or may not be an advantage. Third, mapping tools disrupt the temporal flow of conversations and impose a spatial organization of contents, in particular they may destroy the reply structure which is typical in conversations. Thus there is an issue of measuring to which extent argument based collaboration suffers from the above problems and may actually impede rather than favor effective common ground building.

The following section explains the methodology used to explore whether the advantages or the disadvantages of argumentation prevail in a large scale discussion on a wicked problem when argumentation is the way the deliberations are conducted (**Table 1**).

Table 1 Advantages and Disadvantages of Argumentation in Comparison with other Online Tools

Advantages	**Disadvantages**
Navigability	High level of formalism steepening the learning curve
Lower redundancy	Reduction in communication bandwidth
Lower conflict	Disruption of temporal flow
Argument map as a facilitator knowledge object	

I.3) METHODOLOGY

In order to explore the development of common understanding, if there is any, data from an experiment held at the University of Naples Federico II in December 2007 in cooperation with the MIT Center for Collective Intelligence. An argumentation-based platform, hereafter called Deliberatorium is used for deliberations on a wicked problem of biofuels, by a community of around 160 master's students. In order to give an initial boost to the discussions, 4 issues are provided to the community, namely "how can the obstacles against the diffusion of biofuels can be reduced in Italy", "what impacts will biofuels have on the economy", "what impacts will biofuels have on the environment" and "do biofuels have a future in Italy". While biofuels are initially thought to be reducing greenhouse gas emissions, according to some, the cultivation of crops used in their production emits more than usual. Although alternative sources suggested such as algae, their economic and technical feasibilities are still in question. As can be seen at this example, the future of biofuels in Italy – like any other wicked problem - has an unexplored

solution space where new solutions suggested can raise new questions, creating a network of sub-problems.

Using this experiment the basis and accumulativeness principles of common ground theory were are observed as in the following:

1. *Basis*: Deliberating on a wicked problem may entail the discussion of multiple topics given the possible involvement of various disciplines. Therefore having *a basis* on every one of these topics may be more challenging for deliberations on wicked problems. In the biofuels example, discussions may spread over topics ranging from politics to technology. In order to observe the range of topics and the intensity of discussion on each topic, a content analysis based on statistical clustering is conducted. In order to limit the degree of subjectivity, typical of cluster analyses, the following steps are implemented:

    a. All words used in the discussions are taken as an initial set of keywords for clustering. The set is refined by the elimination of stop words, i.e. very common words such as the conjugations of the verb "to be" in Italian.

    b. In order to take different forms of a word such as adjective, adverb etc. as a single entity, each word is assigned to its stem. Therefore the set of keywords are composed of stems rather than words after this step. Stems are identified through Porter's stemming algorithm (Porter, 2006) which has been implemented using Python.

    c. The set of terms is further refined to obtain more meaningful clusters. Some words may be so common that they do not depend on the topic in question. Therefore the ones with higher discriminative quality are retained through vector space modeling (Salton, Wong, & Yang, 1975).

d. All terms are clustered using a network based algorithm. The network is composed of terms where terms correspond to the nodes. Links between terms are formed if they are used within the same post. The network created by this way is then decomposed into its major clusters using the algorithm of Clauset et al. (2004).

e. Clusters obtained at stage d yield clusters of terms which are assumed to be representative of different topics of discussion. However a post can contain terms associated with different clusters. Therefore number of terms associated with every cluster is counted and the post is assigned to the cluster which is represented through a higher number of terms.

2. *Accumulativeness*: One can consider consensus as the departure point for common ground building. In a group where there is real collaboration, the consensus is supposed to be developed with contributions highlighting the accumulative characteristic of the common ground theory. Therefore common ground building is a process rather than a stationary event. Longitudinal analysis is used in order to verify the accumulativeness of common ground. This permits to see whether a contribution made on a certain topic is preceded by a period in which the author explored the contributions of the community in the same field.

I.4) EXPERIMENT AND DATA

The aforementioned Deliberatorium that provided data for the analysis is a medium where the deliberations take place in a collective manner. Argumentation method is applied through a process in which issues and ideas submitted by users are supported and attacked by arguments from other users. This in turn creates a chain of user posts that reveals the pros and cons behind each idea.

A critical issue is the users' argument mapping skills. Since participation is open and voluntary, a sufficient number of skilled argument mappers may not be available or users may not spend enough effort to learn to use the formalism properly. Thus a moderators' corps was introduced. Moderators help ensure that posts follow the basic rules of structured argumentation and are properly located. They are expected to provide feedback to users to help them learn how to use the tool correctly. Consequently, in the Deliberatorium virtual community, there were three types of user: moderators, authors and readers/voters.

Moderators were charged with the usual tasks of filtering out redundant and irrelevant posts. They were also in charge with ensuring that the argument map was well-structured, i.e. all posts were properly divided into individual and non-redundant issues, ideas, and arguments, and were located in the relevant branch of the argument map. This involved classifying and sometimes editing posts, offering suggestions to authors, aggregating similar arguments, and occasionally re-organizing the overall argument map so that related topics are grouped into the same branch. A team of 4 student moderators was selected and trained in argument mapping before the test. While moderators can be selected in a bottom-up way, they were selected beforehand because of the limited time for the experiment not allowing observing the emergence of a self-organized hierarchy. The on-line argumentation process developed as follows:

1. Authors posted and edited issues, ideas, and pro/con arguments and produced an argument map similar to that in **Figure 1**. To help user to structure their contributions, arguments were posted using an on-line form. The form was designed to facilitate unbundling contributions into issues, ideas and arguments. Nobody, except moderators, could edit a post authored by someone else.

2. All users (including moderators, authors and readers) could rate arguments and ideas, as well as send comments to authors through threaded discussion forums associated, like wiki talk pages, with each post. Rating was anonymous.

3. Posts were initially given a status of "pending", and could only be certified by moderators. Until a post was certified, it could not be rated and nobody, except its author, could link any other posts to it. The rule was that only certified posts would have appeared in the final, publicly available, version of the argument map and it was not possible to attach additional posts to uncertified posts. This helped to limit useless proliferation that can be observed in discussion forums in which even bad or deliberately provocative posts can trigger long but usually low value-added discussion threads.

Deliberations conducted through Deliberatorium are recorded in a database. Events conducted by users such as reading, editing, voting or writing a post can be followed through time. The database holds about 110000 time-stamped events along with the content of the deliberation.

I.5) RESULTS

Clustering of user posts yielded clusters around three major axes that can be roughly stated as economic, political and environmental. **Table 2** provides an example of the terms associated with different clusters.

Table 2 Exemplary stems and the related words that are used during discussions

Cluster no	Stems-Words
1	malnutrizion (malnutrizione) monocultur, (monocultura) parlamentar, (parlamentari) autoprod, (autoproduzione) profitt (profitto, profitti)

2	multinazional (multinazionale, multinazionali) nazional (nazionale, nazionalità, nazionali, nazionalismo) liberalizz (liberalizzare, liberalizzato, liberalizzazione) industr (industria, industrie, industri) concorrenzial, (concorrenziale, concorrenziali)
3	atmosfer (atmosfera, atmosferiche, atmosferica, atmosferico) anidrid (anidride) bioetanol (bioetanolo) bioenerget (bioenergetiche)

These clusters constituted the departure point for assigning every user post to the most related cluster. Posts are assigned to the cluster that is represented through a higher number of terms within that post. While the presence of diverse topics provided support for the existence of the *basis* condition of common ground, there is also need to verify the other condition, i.e. *accumulativeness*.

For this purpose a longitudinal analysis is done using constant coefficients model. Longitudinal analysis is appropriate to use with panel data where observations are made at consecutive times. The results suggest that the number of posts that a user reads on a certain topic in the past leads to an increase in the number of posts he/she authors on the same topic.

I.6) IMPLICATIONS

Implications can be regrouped as theoretical and methodological. The former entails the use of common ground theory as a basis for evaluation of an online deliberation tool. This also constitutes a reference point for future evaluations on other media. Methodological contribution has been in the way the common ground was assessed. Longitudinal analyses have been used in various fields particularly in sociological fields. Its use as an assessment of common ground building complements the methodological contribution made here. Using this approach one can compare different tools as well. A significant common ground building activity through

argumentation will therefore be a motivation to make comparisons with alternative tools. The study also gives an idea on the extent to which diverse aspects can be explored by a student community. A satisfactory outcome would be an encouragement to conduct similar experiments also with expert communities.

Last but not least, this study can be considered as an initial step in the development of tools that aim at leveraging the wisdom of different people in different locations. This is an ambitious but necessary target in the face of problems requiring urgent solution such as the unequal distribution of wealth and the global warming.

I.7) LIMITATIONS

First limitation sources from the subjectivity of the clustering method. As Hair et al. (1995) also mention; clustering is the most subjective statistical method, requiring caution in interpretations. Second, a similar problem arises from the stemming algorithm. Although stemming can group words sourcing from the same root, this does not necessarily correspond to 100% accuracy. One word may have different meanings in different contexts. For instance the verb "cry" may mean "to shed tears" as well as "to shout" (Dictionary.com). The algorithm may also cause over-stemming, i.e. the unnecessary removal of letters because of being considered in a suffix. While the words "provable" and "probable" can yield respectively "prov" and "prob" after stemming, they both origin from the Latin word "probare", "to prove or to test", and the idea of testing connects the meanings of the words (Porter, 2001). However the meanings are not parallel. "Provable" means 'able to be proved'; probable does not mean 'able to be probed'. Removing -able from probable is an example of over-stemming. Another limitation of the stemming algorithm is the handling of synonyms which would be taken as totally different words although

the meanings are the same. However these limitations are not unavoidable. They can be alleviated by manipulating the Python code that is used to apply the stemming algorithm as well as by carefully checking and manually correcting the results of stemming and clustering steps.

CHAPTER I. PROBLEM SOLVING IN THE INFORMATION AGE

Many call the current age “The Information Age” characterized by the rapid flow of information across the globe following technological innovations such as the Internet. The problems that the humanity currently faces have become much more complex in comparison with the challenges of the last century. In order to see how this complexity can be dealt with there is need to explore the shift in the nature of problems and the accompanying approaches within the last century. This chapter gives the development of views in a chronological order along with the advancements that led them. Finally the necessity of collective approaches is discussed as a way to overcome the problems of Information Age.

I) A NEW PARADIGM OF PROBLEM SOLVING

I.1) MECHANISTIC APPROACH

The Industrial Revolution introduced the substitution of men by machines. Men were no longer in charge of the whole production process, they were neither autonomous in conducting their tasks. Increasing mechanization went hand in hand with the standardization of tasks which had been transformed into simple elements. Thus they could be repeated by different people in pretty much the same, standard way. Ackoff (1973) explains the major view of this age through two major characteristics:

1. Reductionism: Everything can be disintegrated into smaller, simpler, indivisible parts such as atoms in physics or cells in biology. Reductionism paves the way for analytical thinking. It stands for providing explanations for phenomena by explaining the

independent parts that they are made of and by gathering these explanations to elucidate the whole.

2. Mechanism: Everything can be explained through a cause-effect relationship as a closed-system. The effect of the environment is therefore neglected. This approach is embodied in experiments held in laboratory conditions isolated from the outside world.

According to Ackoff, reductionism and mechanism properties render the prevailing view of the Industrial Age deterministic and mechanistic. The former stands for obtaining the same effects from the same causes under identical, controlled conditions while the latter represents the insignificance of teleological explanations that mediate the cause and effect such as functions, goals and purposes.

This view was first put into practice by Scientific Management, also called Taylorism (F. W. Taylor, 1914) that advocated the use of empirical methods to reduce inefficiency. Before then, at the outset of the 20th century, productivity was not a concern for manufacturers. There was not more or less productivity but working hours that made the difference in the quantity. However scientific management deeply changed this view. Taylor analyzed the tasks and decomposed it into its constituents. This revealed that a considerable amount of manual work was unnecessary changing the traditional view dramatically that used to glorify craftsmen. Then Taylor redesigned the remaining work in the simplest way possible putting the least mental and physical strain on labor. Finally these redesigned parts were put together to be followed by the whole labor force. Taylorism had such a huge impact that the productivity movement boosted economic growth. Although many other approaches rose within the last century, their origin has always been Taylorism. Industrial Engineering, the assembly line of Henry Ford, division of labor, job rotation, kaizen and lean production, they all bore the task analysis principle (Drucker, 2000).

I.2) SYSTEMS APPROACH

Systems approach developed as a reaction to the reductionist view of the mechanistic approach that prevailed before. It worked well for explaining phenomena that emerged as a result of a limited number of isolated causal chains. But events that entailed several parameters remained unexplained.

Von Bertalanffy (1972) criticized the two principle epitomes of the classical science, namely systems as automata and systems as products of chance. The former is based on Descartes' suggestion that the world works like a machine, as clockwork. The later was expressed by Darwin's natural selection idea. Both thoughts were very successful. For instance an organism was made of organs that were composed of cells and cells that were composed of organelles that work in perfect harmony. In a similar vein the occurrence of random events explained a wide range of facts through the evolution theory. However Descartes had used his epitome as an explanation for the existence of God. In the end the order at system level had a spiritual root while the natural selection idea was self contradictory being unable to explain men's own innovations which did not happen to exist out of accidental, chance-driven events.

The systems approach existed even in ancient times as Aristotle stated "whole is more than the sum of its parts". However this statement was acknowledged only at a philosophical level being left out of science due to the incapability of mechanistic approach in explaining the complexity in relations among the components of a system. In a parallel but disconnected way, cybernetics development was boosted by research in military fields. Therefore the Second World War had an accelerating effect on the systems theory.

Systems approach was embodied in fields such as Operations Research (OR) and policy sciences. The systems approach of 1950's and 60's targeted optimizing the behavior of systems and keeping them in that state (Checkland, 1985). For this purpose efficiency was the key term. For instance the objective in an OR problem is finding the most effective solution which is expressed in terms of efficiency by Ackoff (1956) as in the problem solving steps below:

1. Measure of efficiency related to the objective is determined.
2. In case the units of different measures are not commensurate they are transformed adequately.
3. The most effective is defined such as maximum profit or minimum waiting time.

Therefore the focus of systems approach has been, attaining a desired goal. However this mostly requires the abstraction of the real world situation implying a trade-off between losing the representative power of the model used in problem solving and the benefits of rendering a concept simpler in order to be able to solve the problem.

I.3) SOFT SYSTEMS THINKING AND COMPLEX PROBLEMS

Checkland (1985) calls the systems approach adopted in 1950's and 60's as hard systems thinking which was characterized by goal seeking behavior in problem solving efforts. However he argues that hard systems thinking is not adequate due to the changing nature of the problems that mankind has been facing. Problems are now embedded in much richer contexts. Therefore simplifications required for hard systems thinking turn out to be oversimplifications of contexts that may be a part of the problem. Soft systems thinking on the other hand, abandons the goal seeking behavior positing that before identifying the goal there is need to define the problem first whereas goal seeking aims at regulating activities among parts of a problem. Thus it assumes a

stationary relationship among those parts through time. However many problems provide with various possible courses of action and implementing a course of action generates new others. The nature of these problems has been identified through various attributes such as being ill-defined (Wilson, 2001), complex (Bar-Yam, 2002) and wicked (Rittel & Weber, 1973). "Wicked" and "complex" problems are hereafter used interchangeably.

Rosenhead & Mingers (2001) posit that two factors are key in explaining the increasing consciousness on wicked problems. First the *complexity* of relations within organizations and among individuals resembles a network rather than dyads or little groups. Second there is *uncertainty* on the impact of every node's action on other nodes' reactions while it is also difficult to understand the dynamics of relations within the network. Even a full comprehension of the dynamics is not enough given the rapid change in context which may change the objectives simultaneously.

Rittel & Weber (1973) distinguish problems as wicked versus tame problems. They claim that scientists and engineers have traditionally focused on tame problems whereas professionalism that rose during industrial age was not enough to deal with societal problems. Social science applications such as city planning and policy planning did not prove to be successful because they intended to mimic natural sciences. This is related to the different characters of problems that they deal with. Natural sciences study definable problems with findable solutions whereas social problems are never solved but solved over and over again. For instance we know that there exists a solution in a mathematical equation or we know we can find what compounds are formed in the end of a chemical reaction. However even for global warming there is no consensus yet. For instance while greenhouse gas emissions' influence is widely acknowledged,

RAND, the name derived from Research and Development is founded during World War II with a military focus. It was initially a division of Douglas Aircraft Company of Santa Monica in Califirnia. Today it is a nonprofit organization funded by the US government as well as the private sector.

One of the contributions of RAND has been the systems analysis, a term suggested by a RAND researcher named Ed Paxson. The interdisciplinary approach adopted by the institution is reflected also in the design of RAND's Santa Monica headquarters. The layout was planned in a way that would enable people from different fields to encounter in their day to day activities. Simplex Algorithm is another innovation conceived by a RAND employee. It was developed by George Dantzig in late 1940s. Linear programming programs became solvable in a very short time with the aid of computer.

In 1970s RAND shifted its focus to policy analysis in both military and non-military settings. (Campbell, 2004)

there is also the belief that global warming is just an artificial problem created by alarmist scientists (Lindzen, 2006).

The complexity school of thought criticized the modernist view for overemphasizing scientific knowledge which was believed to be enough to give a good enough understanding on the consequences of actions. However real world situations were not closed systems taking place in conditions under control as assumed by RAND Corporation applications or operations research approach (Rosenhead & Mingers, 2001). In an open system, the context has implications on the so called 'cause' that can have diverse results in different situations. Therefore the concern is to decide on where the problem lies in an open system and from where to start to solve it within the interconnected network of open systems. In the light of these propositions Rittel & Weber (1973) provide ten distinguishing properties between tame and wicked problems:

- Formulating a wicked problem is a problem in its own right. The search for the necessary information to solve the problem requires an idea on possible paths of solution. For instance what would the question on global warming entail? Is it a question of pollution? If so, is it because of the lack of necessary regulations or the lack of control mechanisms? For the latter, why don't the control mechanisms work? Is there a corruption or an authority problem? Answering each of these questions requires another set of information. Defining the problem is therefore, somehow, the same as finding the solution.
- Wicked problems require exploring a network of problems, solutions, actions. Fully exploring such network is impossible due to the constraints such as time, finance, expertise etc.
- Judgments on wicked problems cannot tell whether the suggested solution is true or false unlike the solutions for tame problems. Judgments of parties differ according to their interests. Given the varying levels of expertise in diverse fields, people cannot come up with the same set of criteria for assessing a solution.
- Testing a solution to a wicked problem is impossible given the open systems that they are embedded in. Replicating the real world situation in closed, under control conditions is impossible. Therefore the only way to observe the results of a solution idea is actually applying it.
- While simulating a wicked problem situation observing the effects of an idea and then undoing it is impossible. Take the diversion of water from Amu-Darya and Syr-Darya rivers for irrigating cotton fields. What once seemed reasonable to increase economic production has had irreversible effects. The world's fourth biggest inner land sea, the Aral Lake shrank by 75% having economic consequences such as destroying fishery. What was once sea is

now a desert and the pesticides used for agriculture are spurred around by wind threatening public health by increasing respiratory diseases. The World Bank decided to finance an $85million project to stop desiccation. However even this figure will be used for a project aiming at saving only Small Aral i.e. one of the two smaller lakes that appeared after desiccation (Pala, 2005) whereas political dispute of neighboring countries is just another aspect of the problem (Nature, 2003).

- Exploring the whole set of solutions for a wicked problem is impossible. The decision to enlarge a set of solutions to assess and to follow one of them is a subjective task. Depending on the answers given to the questions in the first property above, the extent of the solution space to discover will change accordingly. In the presence of such diversity, selecting the right course of action is important.
- A solution that may work for a wicked problem does not apply for a similar wicked problem. In spite of similarities, open systems imply context-driven effects on the problem that are peculiar to every situation.
- Wicked problems can be considered as interconnected problems at various levels. The removal of the cause reveals the problem from which the original problem resulted. Thus solving the problem at a higher level would have a wider effect. However coming up with an acceptable solution for a higher level of problem is more difficult as well. For instance illegal immigration can be considered the result of insufficient number of patrol troops at country frontiers. Alternatively it can be a result of the imbalanced distribution of wealth across countries. The latter stands for a problem at a higher level and requires much more effort for a solution.

- Divergence in the expected results of a solution suggested for a wicked problem can be explained in various ways. The explanation depends on the approach taken in the solution. For instance tax deductions can be proposed to stimulate economy in a crisis. In case the effect turns out to be less than expected, the advocates of tax deduction can still claim that without tax incentives the situation could have been worse.
- As also implied by the irreversibility of actions taken in solving a wicked problem (5^{th} property), the solution provider's mistake is not tolerated as it could be in a rational scientific context. In scientific community, an idea is tested through hypotheses. A hypothesis is considered valid as long as it is not rejected . However once it is rejected it does not make the owner of the hypothesis blameworthy.

Various problems that occupy the international agenda today such as imbalanced wealth distribution, immigration, human development, global warming, desertification etc. are wicked. Wicked problems apply also to organizations. Taylorism was responding to the needs of the tame problems of the manual manufacturing world. However the nature of business is much different today in presence of much wider sources of energy and technology. The job market is now being dominated by workers who are embedded in interpersonal exchanges, solving complex problems, every time in indifferent contexts. While this type of workers is growing at a rate of 4.2% in USA, traditional production workers increase at 3.0% and transaction workers, i.e. performers of routine interpersonal exchanges such as secretarial staff, grow at a rate of 1.8% (Manyika, Sprague, & Yee, 2009).

The omnipresence of wicked problems has had consequences on traditional disciplines as well. For example, Industrial Engineering focused on increasing efficiency in production systems in the first half of the 20^{th} century through notorious time and motion studies. Following the World

War II attention was given also to analytical techniques as could be observed in the rise of Operations Research. Human-machine interaction became a new area studied in laboratory conditions. However the working environment changed since then. There is now growing interest in cognitional aspects of work while ergonomics remained within the confines of human-machine interaction. Adoption of Information Technologies has changed the way tasks are done However in such dynamic context, the ideas of ergonomics obtained in controlled laboratory conditions is of limited use. With the new rules of the game, Industrial Engineering is suggested to rediscover the importance of empiricism to complement widely used mathematical and laboratory studies. There is better need to comprehend the dynamics of an open system. Therefore for industrial engineers there is need to be more knowledgeable about social sciences which has various implications for the industrial world such as social networks, power relations, organizational culture, social conflict (Bailey & Barley, 2005).

II) Collective Intelligence for Solving Wicked Problems

In his evolution theory Van Valen (1973) published the taxonomic survivorship curves for certain species to show their extinction rates. In his graphs he plotted the number of species in a family or genera against the duration of member species' fossil record. While such study had been done before, Van Valen plotted the logarithms of the number of species rather than a simple count of them. This revealed that survivorship of species tend to be log linear. Log linearity implies that the probability that species will extinct does not depend on its record of evolution. In other words, they do not become any better in avoiding extinction in the course of evolution. Species decay at an exponential rate with a constant proportion of survivors going extinct at every time interval. He explained this as an arms race between coevolving populations with the goal of staying ahead of the other, a phenomenon called the Red Queen effect. This expression

comes from what the red queen said to Alice in Wonderland: “here, you see, it takes all the running you can do to keep in the same place” This implies that between two competing species undergoing an antagonistic coevolution, one will always be in disadvantage as the other develops new adaptations. In such situation the lagging species will either develop another adaptation or will go extinct through natural selection. Coevolution will therefore lead to a log linear survivorship if the deterioration resulting from competing adaptations has a constant pace.

The Red Queen Effect idea has been adopted by scholars to explain phenomena such as the competitiveness of organizations (Barnett, 2008) and the struggle of antivirus companies against computer virus writers or Internet service providers against spammers (Carlsson & Jacobsson, 2005). By the same token, Bar-Yam (2002) considers the modern organization and the environment that it is embedded in, as coevolving entities. However the environment has changed considerably within the last century. Inter and intra connections in economic, transportation and communication systems spread rapidly, carrying the effects of the Information Age even to our daily lives. Today one may get nervous if an item he/she orders online is not delivered in a couple of days whereas the same delivery – if online ordering could have been possible – a century ago may have never arrived. While placing the order does not require much effort, such rapid delivery is the result of a complex, highly connected network of deliverers, producers, online operators etc.

This pace of change in the environment is here to stay. Therefore according to the coevolution idea of Van Valen, in order to survive, organizations should respond to the increasing complexity surrounding them. Bar-Yam posits that traditional hierarchical structures do not exhibit any complexity although they constitute a collective action. In hierarchical organizations, at each hierarchical level lower parts are coordinated and controllers at each level are responsible for

transferring the policies of the upper level. Therefore the actions a controller take must be in line with the actions in the rest of the system. Thus the whole hierarchical organization is embodied in one single person, i.e. the person at the highest level. However an organization whose collective action is under the control of a single person cannot behave more complex than that person individually can. On the other hand adaptation to a complex environment requires additional effort in defining the problem, exploring possible solutions and integrating expertise from different fields. Thus it requires knowledge that cannot be borne by a single person (Tsoukas, 1996) implying a need for collective approaches in problem solving. This approach is embedded in matrix organizations that rose as an alternative to unwieldy hierarchical ones.

Changing nature of the business world also introduced what Drucker (2000) called *knowledge workers*. Unskilled employee of the Industrial Age could be replaced by another one easily as a result of the Taylorist approach. Knowledge workers however have an egalitarian relationship with the employer because unlike manual work employees, they own the means of work. The knowledge they have is completely portable whereas the traditional employee cannot take away his/her job shop (Drucker, 2000). Therefore the traditional boss-subordinate relationship leaves its place to a new management approach. The non-repetitive tasks of knowledge workers require high cognitive effort. Executives should be good planners and coordinators as well as technically knowledgeable to be able to understand knowledge workers.

Therefore the rise of wicked problems and knowledge workers came hand in hand. As its definition implies, knowledge work requires working with others because knowledge workers collaborate and interact to solve wicked problems in changing contexts (Manyika et al., 2009). Handling wicked problems in groups makes sense mainly due to the four of the aforementioned characters of these problems. First the problem has a network structure that is impractical to fully

discover given the limitations on time, financing and expertise. Second the solution space is boundless. Therefore the larger is the group of problem solvers, wider will be the solution space explored. Third wicked problems can touch upon various expertise fields requiring assistance from experts of related fields. Fourth not only the experts but also the stakeholders that have a say on the question in hand should be integrated to the problem solving process. Having their consent is important given the irreversibility of actions taken. Working in groups is convenient also in order not to overburden an individual who is very unlikely to be an expert in all related fields.

Therefore integrating different perspectives is a necessity to solve wicked problems. While gathering people from such diverse fields may be a challenge, it can be an opportunity as well, provided that the collective intelligence can be stimulated. Collective Intelligence is based on the idea that a group of individuals can come up with a solution to a problem in a way better than any of its members can (Heylighen, 1999). As Raymond (2001) puts it, "given enough eyeballs, all bugs are shallow". (S. E. Page, 2008) posits that diverse skills and competences can allow a richer problem analysis and solution while Vennix (1996) shows that differences of views can have productive consequences. Difference in functional backgrounds contributes in developing a better understanding on the problem and the generation of ideas (Pelled et al., 1999). However group practices have not always been successful (Barron, 2003). There are situations where the diversity of knowledge can actually hinder effectiveness (Cohen & Levinthal, 1990) and some groups may overemphasis cohesion and neglect critical thinking, hampering a fair analysis of alternative solutions (Janis, 1982).

Sunstein (2006) delineates four major causes for the failure of groups in making accurate, truthful and reliable decisions:

- *Error Amplification*: Social dynamics mostly work in favor of the group error risk. As Condorcet Theory applies, with an increasing group size, making an error as a group approaches 100% probability if each individual is more likely to be wrong.
- *Hidden profiles*: Groups may not be able to come up with the necessary outcome even if they hold all the necessary information due to the overemphasis of the commonly shared views across the group. Information held by all members can have more influence in discussions simply because they may be mentioned more often during deliberations or social issues such as reluctance of lower status people may impede information disclosure.
- *Cascade effects*: When people reveal their ideas in a row, if the majority of the people who revealed ideas initially support a certain view, those who follow would feel pressure to make the same choice. This urge would be much higher in case the follower is indecisive. A similar effect can be observed due to the reputational concerns when people support a view not because they believe in it but just because they do not want to look foolish in front of others. The result would therefore be a premature and fake convergence of ideas
- *Group polarization*: When members have a preexisting inclination towards a certain view before the deliberation, people may end up with a position more extreme than the average opinion. Social influences may take a role as well. People may change their views in the direction of what the others think e.g. to please them or to be accepted. Finally overconfidence that may rise as a result of being supported by the others can reinforce extremist views as well.

The following two points however are favorable for letting groups outperform its members:

- When there is a certain level of support for the correct solution at the outset of discussions, the advocates of this solution can convince the rest of the group.
- The correct solution is widely supported when the question in hand is a “eureka” question, one that has a clear answer such as the distance between Naples and Rome. In such cases, the presence of even a limited number of people who know the answer is enough to convince the group.

Although the aforementioned failure causes outline the major problems encountered in experiments held with small-scale, physically co-located groups of individuals they can apply for large scale discussions as well. Going back to the ongoing efforts for saving the Aral Lake, one can make the following non-exhaustive list of interested parties and experts that may be suggested to take role in the problem solving process:

- Governments of countries neighboring the lake
- Governments of the upstream countries in Amu-Darya and Syr-Darya basins using rivers for hydro power generation
- Financers such as The World Bank
- Scientists
- Environmentalists
- Representatives of people who need irrigation for cotton fields
- Representatives of people living around the lake and who seek for a livelihood.

Finding a solution to such a problem with the involvement of a big community might seem impossible. It is true that the complexity of wicked problems can be discouraging. However the Information Age is not only associated with increasing complexity but also with the rise of

collective intelligence. Along with the emergence of new communication technologies, namely the Internet, collective intelligence has been characterized increasingly by interactivity, user generated content and openness. Thus online tools can be used in leveraging the collective intelligence of large groups, Chapter II will put emphasis on different features of available tools explaining their weaknesses and strengths.

CHAPTER II PICKING UP THE RIGHT ONLINE TOOL AS A MEDIATOR OF LARGE SCALE ONLINE WICKED PROBLEM SOLVING

Chapter I is concluded by positing that large scale deliberation can be leveraged to deal with wicked problems and online tools can be used to mediate the problem solving activity of the community. In the Information Age, Internet would be a straightforward answer enabling a cheap way of gathering geographically dispersed people. However the ways people can gather is innumerable given the large scale of products ranging from emailing to online Delphi method applications. Therefore first current online tools are described through two major classes, namely Group Decision Support Systems and Website-based tools. Second different dimensions of online tools are given, providing a list of aspects that vary across different tools. Finally they are assessed based on the requirements of a wicked problem solving task.

I) GROUP DECISION SUPPORT SYSTEMS (GDSS)

There are ambiguous classifications made on online tools. Computer-Mediated Communication (CMC), Computer-Supported Cooperative Work (CSCW), Group Support Systems (GSS), GDSS and Electronic Meeting Systems (EMS) all have definitions overlapping to a certain extent. Dennis et al. (1988) posit that GDSS are more task-oriented aiming at facilitating decision making or solving a problem whereas CSCW tools are designed for better communication across the members of a small group as in the critiquing of a document. Chidambaram & Jones (1993) use GSS-GDSS and CSCW- CMC interchangeably. However both researches see EMS as a way of combining the task and communication oriented focuses to support group work.

This ambiguity is understandable given the wide range of tasks that can be mediated by online tools supporting group work. These tasks cover, among others, communication, idea generation, planning, problem solving, negotiation, conflict resolution, analysis and design (Dennis et al., 1988; Huber, 1984). To complicate things further, not every group work is supposed to support every task. Therefore different tools are adapted for different purposes and most of the time they are designed for specific group works rendering it difficult to draw general implications. Thus the best generalization possible is the common use of hardware, software and verbal components by all of them.

Huber (1984) makes a general classification of information sharing and information use features. First use of numeric information enables producing information by testing different scenarios in real time. This is helpful in order to prevent divergence. Second presenting relational information is another support that GDSS can provide. PERT charts, critical paths or decision trees are examples that existed before computers and with the rise of the digital world they are applied in GDSS tools as well. Use of relational information lets compare different alternatives, an important precondition for decision making.

GDSS tools can support traditional methods used for decision making that existed prior to the emergence of IT. For instance Delphi (Fussell, Kraut, & Siegel, 2000; Kesten Green et al., 2007; Linstone & Turoff, S.d.) and Nominal Group Techniques (Delbecq & Van de Ven, 1971) comprise of steps that can be applied online.

II) WEBSITE-BASED TOOLS

This group involves the oldest forms of online communication such as emailing as well as more recent tools such as wikis. Although they are considerably basic compared to GDSS, they had an

undeniable role in globally boosting information sharing, creation and accumulation. Therefore the business world has been inspired by success stories such as Wikipedia and Linux (Bonaccorsi & Rossi, 2003; von Hippel & Katz, 2002; von Krogh & von Hippel, 2006; Raymond, 2001).

As for GDSS, website-based tools cover a wide spectrum of products as well. Klein et al. (2007) classify these tools in three groups:

II.1) SHARING TOOLS

Internet owes a great deal of its fame as a collaborative platform to this category of tools. Emailing that existed even before Internet already allowed collaboration to a limited extent. Usenet paved the way for the development of forums. Different from forums, Usenet is distributed over a network of servers where each server forwards messages to another. With the launch of World Wide Web, forums had widespread use by attracting people who were not technically savvy. In addition to the timestamps and archives of Usenet, they integrated features such as color, voice and hypertext to their design (Rafaeli & Raban, 2005).

Blogs have become very common letting people publish content continuously. This trend led to groups of blogs such as Slashdot.org creating online communities around them. As individual blogs may let bloggers gain reputation, group blogs make use of reputation tools in a similar manner e.g. each Slashdot community member is assigned a “karma” value which is determined by the community’s evaluation of individuals’ contributions. Same logic applies to online recommendation sites such as answers.yahoo.com where those who provide answers have ratings as a proxy of reliability.

Wikis have had widespread use particularly with the rise of Wikipedia. Wikis let people edit pages using a web browser. Anyone can edit anything. Their use for creating dictionaries and encyclopedias has proved to be a great success.

In online social networking people get in touch with others through personal acquaintances transferred to the online medium. Its use has spread with the emergence of sites such as orkut, facebook and LinkedIn. A variation of online social networks inheriting the informative aspect of blogging is Twitter. It has been an efficient way of surpassing the censure that the Iranian Government applied during the manifestations against the elections results in 2009.

Twitter started as a start up project in 2006. "Twitterers" can post what they are doing, in a message limited to 140 characters by using texting, instant messaging or the web. Once a twitterer sends the post, all subscribers who follow him/her receive updates.

While this could be a cool way of staying in touch for teenagers, it also proved to be a new way of communication after the 2009 elections in Iran. Supporters of the opposition candidate Mir-Hossein Mousavi protested the results of the elections that were in favor of Ahmedinejad. However the censorship on media prevented the information flow from the country. Iranians responded by maintaining the information flow through twitter. For this reason the US State Department asked Twitter to postpone the regular network upgrade.

Twitter is free, mobile and fast. By using hash tags (#) users can tag their posts facilitating their search and retransmission by other users. What makes Twitter suitable for a mass protest is its openness unlike facebook or emailing, letting everyone to see and hear what is going on. Source: (Grossman, 2009)

Rafaeli & Raban (2005) suggest that information may be regarded as a public good which has three major characteristics according to the Penguin Dictionary of Economics. First their consumption by one does not reduce their availability for others. Second they are non excludable, i.e. their consumption cannot be prohibited by anyone. Finally they are non-rejectable, because of their vital importance. Therefore public goods are vulnerable to the free riding problem, i.e. benefiting from a public good while not making any contribution for its provision (Marwell & Ames, 1981). In a system where information is shared directly between people such as in an online forum, limited number of knowledgeable contributors may face a high level of information sharing request discouraging their long term participation. However attaining a critical mass of contributors may alleviate the free riding problem. Markus (1987) states three factors promoting this effect in the creation of public goods in interactive media. First people have heterogeneous interests increasing the likelihood that there will always be someone to make a contribution on a certain task. Second concentration of resources promotes the creation of a public good where dependency across the nodes of the interactive network endorses information sharing. Finally geographical dispersion has a similar effect creating an incentive for those who are more distant to use online media.

Once the critical mass is attained, free riding ceases to be a problem due to the increasing diversity in interests, resources and location. In an online medium the physical absence of free riders render them invisible. Actually free riding is much more preferable to negative contributions which increase redundancy and impede navigability. However the bigger is the share of active contributors in a community, the better will be the interaction. The real challenge sources from the discussion of controversial topics. Having a very loose structure, sharing tools have almost no standard to support a valid reasoning. Therefore the posts are open to bias. Wiki

edit wars is a result of this loose structure where a page goes back and forth between different versions representing different camps on a controversial issue such as abortion, evolution etc. Alternatively it may represent a stubborn insistence on a trivial issue as two people may have a dispute on whether a chocolate structure named “coulage” really existed (Viégas, Wattenberg, & Dave, 2004). Therefore unproductive debates can extend to the issues that are supposed to be uncontroversial as well. A similar example is flaming, i.e. hostile interaction between the online community users that may take place in online forums or email lists. Such hostility can significantly decrease the signal-to-noise ratio in a forum discouraging to stay in the online community.

II.2) FUNNELING TOOLS

Funneling technologies help an online community to reach convergence. E-voting achieves that simply by applying usual election principles to the online medium. It lets users make selection among a number of available options and reveals the most supported option which is what the online community converges on. Another funneling tool, prediction markets, are based on placing bets on the occurrence of a particular event (e.g. election results) or a certain parameter (e.g. sales figures of the next quarter) which works like a futures market. Their accuracy has encouraged private companies such as Google to adopt them (Cowgill, 2005). A notorious example is the Iowa Electronic Markets, a nonprofit prediction market that had accurate estimates for the American election results. However both e-voting and prediction markets do not let users provide alternatives among which they are supposed to make a selection. The alternatives are given prior to the interactions

Founded in 1996, Hollywood Stock Exchange is a prediction market working as a virtual stock market where users can buy shares of celebrities and movies. Prices rise with a blockbuster's opening or freefall with a disappointment in the box office. People sign up for the stock exchange (HSX) and create a portfolio of "investments" with their "Hollywood Dollars". They can buy stocks in movies or they can invest in StarBonds. A StarBond reflects the success of the movies of an actor or an actress in the box office. It is also possible to bet on short term events such as the Academy Awards. The market had accurate estimates in all these investment types.

The business model of HSX is based on demographic information that users provide while signing up for the market. HSX can collect information on the targeted audience for a movie and sell that infmation to the studios such as MGM and Lions Gate Entertainment. Entertainment companies can create their advertising campaigns in light of this information
Source: (King, 2006).

II.3) ARGUMENTATION TOOLS

Online argumentation tools make use of the argument theory to mediate interactions. Argumentation theory is based on the idea that every opinion can be decomposed into its elements, i.e. the conclusion it makes and the premises that lead to that conclusion. Scholars had different philosophical approaches in argumentation leading to different classifications of elements that make up a statement. While approaches may be somewhat different they all enable decomposing opinions to their constituent elements, increasing the visibility of essential

components. Such classification also enables presenting content in a concise manner while rendering the logic behind more evident, implicitly supporting evidence based reasoning.

According to Toulmin a good argument should be able to justify a claim against criticisms (Toulmin, 2003). In his book "The Uses of Argument" he identified the following interrelated classes for analyzing arguments:

Claim, the conclusion made in an argument

Ground, statements supporting the claim such as facts and expert opinions

Warrants, methods of connection between grounds and claims

Qualifiers, expressions stating the degree of certainty in the claims

Backing, credentials in support of the warrant

Rebuttals, statements that constitute exceptions to the validity of claims

In light of this classification, Toulmin, Rieke, & Janik (1979) represent the statement below as in the following figure:

"George has died without leaving a known will, and the question arises, 'What will happen to his widow, Mary?'"

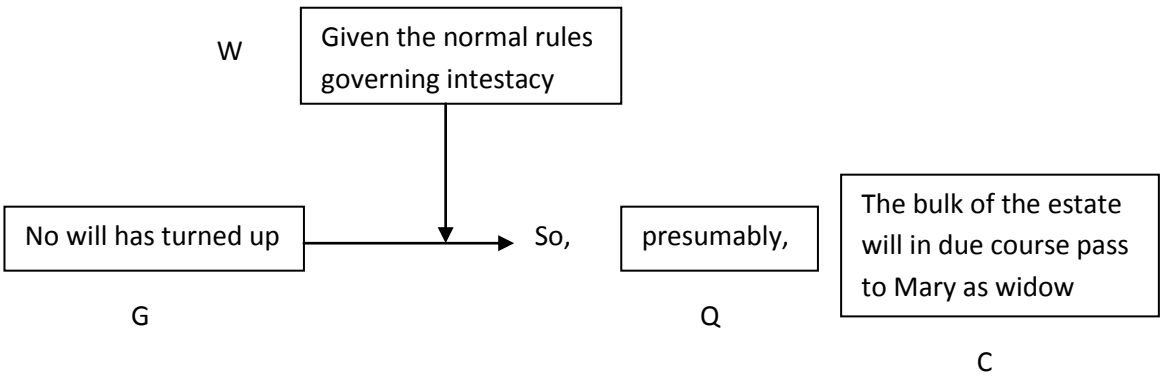

Figure 4 Toulmin's argumentation model (C: Claim, W: Warrant, G: Ground, Q: Qualifier

Walton (1989) provides schemes for assessing the validity of, what he calls plausible arguments. Plausible argument makes claims that are not absolutely true but are valid conclusions based on available premises that represent the information in hand. Walton's schemes provide critical questions to verify premises' validity as in the following:

Table 3Argument scheme examples (Walton, 2006)

Argument Scheme	Argument Structure	Critical Questions
Expert opinion	Ground: a asserts that A is true Claim: A may be plausibly taken as true	Is a in a position to know whether A is true? Is a a trustworthy source? Did a assert that A is true?
Popular opinion	Ground: A is generally accepted as true Claim: There is reason in favor of A	What evidence supports the claim that A is generally accepted as true? Even if A is generally accepted as true, are there any good reasons for doubting it is true?
Causality	Ground: There is a positive correlation between A and B Claim: A causes B	Is there really a correlation between A and B? Is there any reason to think that the correlation is any more than a coincidence? Could there be some third factor, C, that is causing both A and B?

Conklin (2005) suggests the Issue Based Information System (IBIS) with the aim of solving wicked problems. IBIS suggested a three group classification as in the following:

- Issues as questions to be answered
- Ideas as possible answers to a question
- Pro/Con arguments as supportive or unsupportive statements for ideas.

IBIS develops as a tree creating an argument map (see Figure 1 in Introduction) thereby facilitating the exploration of content as well.

The major drawback of argumentation is the steep learning curve that it imposes. This is mainly due to its incompatibility with the traditional way of reasoning that people are used to in daily life. In a conversational discussion, the reasoning follows a temporal order whereas argumentation supports a logical flow.

III) ONLINE TOOL CHARACTERISTICS

Given the wide range of available tools, there is need to classify them in order to better delineate the tasks that they are good at supporting. Even though there is a good deal of research done on available online tools, they have various dimensions that need to be taken into account. Therefore before tackling with wicked problem solving by large scale, online groups; desired tool features should be determined in terms of every dimension.

There are several dimensions that take different values depending on the nature of the online problem solving activity. Table 2 provides a list of dimensions adapted from Dennis et al. (1988).

1. Individual Characteristics: Group members' personalities, expertise, demographics can all have implications for problem solving. Diversity across members can enable exploring different perspectives or create a risk of conflict. While it might be possible to select people with desired characteristics in an organizational setting, it is impossible when the problem solving activity takes place in an online medium, open to public access. In such cases, the tools should be able to highlight the reasoning against tensions that may arise.

2. Group Size: Group size is an important criterion in the tool selection given that not all tools can handle the cognitive load of a large scale discussion. Due to their nature, wicked problems require large groups. With increasing problem complexity, there would be need for expertise from a higher number of fields. However if the level of structure is very high, attracting users gets difficult due to the higher learning effort required to use the tool (see the degree of structure heading below).

3. Past Group Experience: If individuals had another problem solving activity with the same people, they would know each person's attitude and area of expertise. Therefore they might better know where to look for the information sought. They might also save from time they need to learn the tool. However that is possible only to a limited extent in an organization and almost impossible in an online community. In the absence of past experience, the tool design should be simple enough to promote participation.

4. Complexity: Every problem has a certain level of complexity on a scale from tame to wicked problems, the latter being more complex. Whether a problem is wicked or tame is given by the problem itself. However problems can be considered more or less wicked/tame among themselves. For instance a new product design project and global

warming are both wicked problems. However given the larger number of stakeholders and its impact, global warming is more complex. More complex problems are more difficult to comprehend and they create more cognitive load on the participants. Therefore the ability of a tool to prevent redundancy and to facilitate finding the related information gets more important with increasing complexity.

5. Degree of Structure: Structure stands for the extent to which statements are fit into a predefined format which is also called formalism. Formalism coerces the way knowledge is presented and shared to enable better communication (Beers et al., 2005). Formalism is not necessarily an advantage because higher level of formalism means spending more time for learning the tool. This can be an obstacle in increasing the community size. However it also helps to alleviate major concerns in wicked problem activities such as ensuring a conflict-free, efficient communication where information is shared in a way yielding maximum visibility and minimum redundancy. This can be achieved by appropriate structure imposed on the way information is shared with the community. Thus online tools should support the appropriate level of structure for efficient communication.

6. Number of sessions: Problem solving can be held at certain points of time or it can take a continuous process. With increasing group size, conducting synchronous sessions becomes infeasible. While tools such as forums and wikis require a continuous process, some GDSS tools are used throughout multiple sessions. An organizational setting is more appropriate for gathering people in a session which is also possible in an off-line setting. Multiple sessions also imply simultaneity while continuous processing is held in an asynchronous manner.

7. Facilitation: Facilitators are experts of the tool used for problem solving. Therefore they are useful to alleviate the problem of learning mostly accompanied by an increasing structure level. Given that there are various task specific GDSS, groups may be expected to learn to use a new tool at every problem solving activity. A hybrid method giving certain responsibilities to both facilitators and users is also possible. All these different situations require different tool designs.

8. Anonymity: Online identities do not have to be real. While this may create problems in building trust, it also gives people time to reflect thoroughly before making a comment and decreases hesitation in disclosing information with the fear of losing face. In an organizational task, anonymity is not likely even if it is desired. If an open, online tool is to be used, appropriate features such as videoconferencing can be implemented to prevent anonymity.

9. Conflict: Conflict is more likely to happen when the problem discussed is on a controversial issue. However the characteristics of participants such as moral values can be another factor particularly when the identities are anonymous. Tools can prevent this by supporting evidence-based reasoning or by certain control mechanisms.

10. Environment: Depending on the problem, environment can be totally online or it may entail gatherings as in a company. Some virtual groups may become "communities" thanks to the ease of exchanging social cues. Therefore to create team spirit, a tool that allows exchanging social content might be particularly desired.

11. Location: Locations can be as many as the number of members. There might be only one location where discussions are held in a room. Alternatively, subgroups may come

together in separate locations. However the latter two are possible rather for organizational groups focusing on solving a problem related to their company, institute etc.

Table 4 Dimensions of online tool characteristics

Dimension	Values	Depends on…	Values for Wicked Problem Solving
Individual characteristics	**Expert vs. novice**	**Environment**	**Unknown**
	Demographics		
Group size	**Small - Large**	**Complexity, Degree of structure**	**Large**
Past group experience	**Whether people worked in the same group before**	**Environment**	**No former experience**
Complexity	**Tame - Wicked**	**Problem**	**Wicked**
Degree of structure	**Loose vs. High structure**	**Complexity** **Conflict**	**High structure**
Number of sessions	**Sessions at certain points of time vs. continuous process**	**Group size,** **environment**	**Continuous interaction**
Facilitation	**Whether the group work is facilitated**	**Degree of structure**	**No facilitation**
Anonymity	**Anonymous or open identity**	**Environment, location**	**Both anonymous and open identity**

CONFLICT	**LOW VS. HIGH**	**PROBLEM, INDIVIDUAL CHARACTERISTICS, ANONYMITY**	**PROBABILITY OF HIGH CONFLICT**
ENVIRONMENT	**ONLINE VS. PHYSICAL**	**PROBLEM**	**ONLINE**
LOCATION	**SINGLE VS. MULTIPLE**	**ENVIRONMENT**	**MULTIPLE**

IV) EVALUATION OF PRESENT ONLINE TOOLS

Dimensions given in section III can be used for the selection of the right online tool for a given situation. These criteria are used to select the best suited online tool among GDSS, forums, wikis and argumentation for solving wicked problems. While wicked problems can have complexity at varying levels this study takes into account problems such as global warming rather than new product development, i.e. problems with higher complexity.

Continuing from the global warming example, discussing such problems has implications at global scale requiring the participation of policy makers across the globe as well as scientists, environmentalists and businesspeople. Such long list that is claiming to embrace the globe is very likely to involve hundreds of people. Therefore the online tool to be selected should be adequate for groups of as many as 200 people. Such large groups are very unlikely to have had former collaborative experience. Therefore no past experience as a group is expected.

Under certain conditions one can have the means for selecting people with desired characteristics such as the level of expertise or demographics. However such possibility exists more in organizational projects while this work is focused on wicked problem solving at large scale that requires attracting the attention of masses. Therefore selecting people with desired characteristics

is not possible. In such context, the community might be formed by individuals with diverse demographical backgrounds with varying levels of expertise and no past group experience. However forums wikis and funneling technologies manage to gather such communities successfully unless the diversity leads to conflicts. Argumentation may be an option as well. However previous online argumentation applications were limited to small groups. Although The Open Meeting Project is an exception, it was a platform for collecting comments from group members where no significant interaction existed among individuals unlike a wicked problem solving activity (Hurwitz & Mallery, 1995). Therefore funneling technologies, wikis and forums are supportive of the group dimension while there is need for more research on argumentation in this regard.

Problem complexity is particularly important given that the problem considered in this work is of wicked nature whereas the solution of such problems through online methods is not common also because of the incapability of present, widely used tools. GDSS are good as decision making tools. However there is also need for problem formulation as a precursor to decision making. That requires a thorough exploration of related expertise fields. However use of forums, wikis and GDSS tend to create a cacophony with a group of 200 people. Argumentation can handle massive information exchange required for the formulation. It owes this advantage to its high level of structure forcing participants to decompose their opinions into its constituent elements and to provide the relational information for the whole set of elements posted by the community. However the high level of structure employed in argumentation with respect to other tools requires a higher learning effort. It is not certainly known whether argumentation can be a solution in large scale given that attracting a big community is important for the solution of wicked problems whereas with a learning barrier this may be a problem.

The large size of the community has implications on the location, environment, anonymity, facilitation and the number of sessions. Gathering such a wide group forces an online method with open access to public. Therefore there would be users in different locations rendering facilitation infeasible and letting users connect anonymously. Finally, continuous interaction prevails over a discontinuous process held through multiple sessions. However these dimensions are not very distinctive as various online tools allow multiple-location, anonymous, non-facilitated and continuous processes.

The evaluation shows that dimensions such as group size, complexity, anonymity and location are determined by the wicked nature of the problem. Therefore the decision to make for selecting a tool is rather dependent on level of structure and conflict which both suggest the use of argumentation. First argumentation decreases redundant discussions increasing the ease of finding relevant information. Second evidence-based reasoning constitutes a precautionary measure against unproductive conflicts. Finally by presenting relational information through an argument map it facilitates to find relevant information and enables a concise way of communication.

Although there is evidence that argumentation can be the right tool for mediating large scale online deliberation on wicked problems, it is not flawless either. First structured design implies a steeper learning curve with respect to other online tools. Second, the classification of elements does not let transmit everything. While this improves signal to noise ratio, it also reduces the communication of utterances that may be somehow relevant but not directly related. The structure of argumentation considers social cues as irrelevant impeding social communication. However social communication has been found to be influential in creating trust (Chidambaram, 1996; McAllister, 1995) which in turn has consequences on team effectiveness (J. M George,

1990; Holmes & Rempel, 1989; McAllister, 1995). Finally the logical flow of argumentation does not conform to the temporal flow of ideas that people are used to in everyday language. Therefore there is need to have a better idea on whether argumentation can be an effective mediator to let large, online groups to solve wicked problems. Chapter III will look into the ways for making that assessment.

CHAPTER III METHODOLOGY

Chapter II ended with the conclusion that argumentation could be the desired method for mediating large scale online communities' deliberations on wicked problems. However there are concerns mentioned as well raising the need for testing whether argumentation is an effective tool as expected. The absolute success in solving a wicked problem would be achieved by coming up with an acceptable solution as mentioned by (Rittel & Weber, 1973) in Chapter I. Among others, this depends on a healthy communication across the community. However along with the increasing number of participants required for solving wicked problems, this becomes problematic. In this study, the focus is the effectiveness of argumentation in providing efficient communication. It is worth noting that efficient communication does not guarantee a successful online wicked problem solving activity given the presence of other prerequisites such as the participation of knowledgeable individuals, incentives etc. However it is also true that the absence of effective communication does guarantee the failure of a wicked problem solving activity. Thus this work aims at providing with a brick on the wall on the way to explore such activities.

I) COMMON SENSE BUILDING FOR EFFECTIVE COMMUNICATION

For better communication, people need to associate the same meaning for the same pieces of information. When that occurs at the community level there would be a common sense that applies for the whole community. However attaining that state as a community does not happen all of a sudden but requires a process of sense making. Klein, Moon, & R.R. Hoffman (2006) define sense making is a continuous effort to understand the relationships between different pieces of information, their implications leading to predictions for the future state of the affairs.

It covers creativity for passing through different steps of the problem solving activity, curiosity to explore related sources of information, comprehension of explored content, mental modeling and situational awareness. Mental modeling stands for the representation of memory, linking different states of interactions while situation awareness is the knowledge state that is attained, inferences made by using the comprehended knowledge.

In a collaborative setting, sense building is suggested to occur through cycles (Deshpande et al., 2005; G. Klein, Moon, & R. R Hoffman, 2006). At the outset every individual bears a personal framework. Such framework enables one to consider the validity of a hypothesis at an early stage of the sensemaking process. Early adoption of a framework allows faster information exploration and easier revelation of inconsistencies permitting to fix them earlier. Frameworks can serve as data while they can be shaped by available data as well. Therefore frameworks are not frozen. They serve as hypotheses and they are prone to be changed along with the acquisition of further detail and questioning as a result of which they can be modified or replaced with a better one. This feedback cycle is required to ensure that the reductive nature of frameworks does not create any problem. Although they are necessary as a way to grasp the complex interconnections between various facts, it risks neglecting certain others. A simple result is not enough for evaluation. Because it is more important to know how the negative result is produced rather than what that negative result is. Klein et al. (2006) provide the figure below as a representation of the sensemaking process composed of elaboration and reframing feedback cycles:

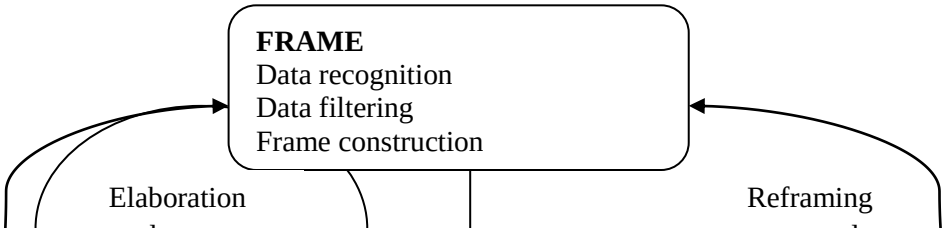

Figure 5 Sensemaking process from Klein et al. (2006)

Sensemaking applies for collective works as well. A collective problem solving activity would be successful to the extent it enables information sharing among participants. However sharing on its own is not enough for building shared understanding. When mutual beliefs, assumptions and knowledge across the community fit into each other, the shared understanding is considered to have yielded to a common ground (Herbert H. Clark, 1996). Any conversation has an effect on common ground leading to its constant update. Throughout interactions, participants seek evidence to see whether they are understood by the community and whether they understood what the community said.

According to Deshpande et al. (2005), at the outset every participant creates an individual framework by selecting what they think is relevant data and detecting reasons to the results with regards to their opinions. This stage can be compared with the elaboration cycle in **Figure 1** above. The set of hypotheses that individuals hold are exchanged through the collective creation of a causal map. During the interactive deliberation on individuals' hypotheses, participants explain the situation in light of context and merge redundant ideas. This in turn leads to the

externalization of the causal chains between issues and the reasons behind them. At the final stage the collective effort of building a causal map pushes participants to question their present frameworks and reframe if necessary as in **Figure 1**. This in the end leads back to the elaboration step (see **Figure 2**).

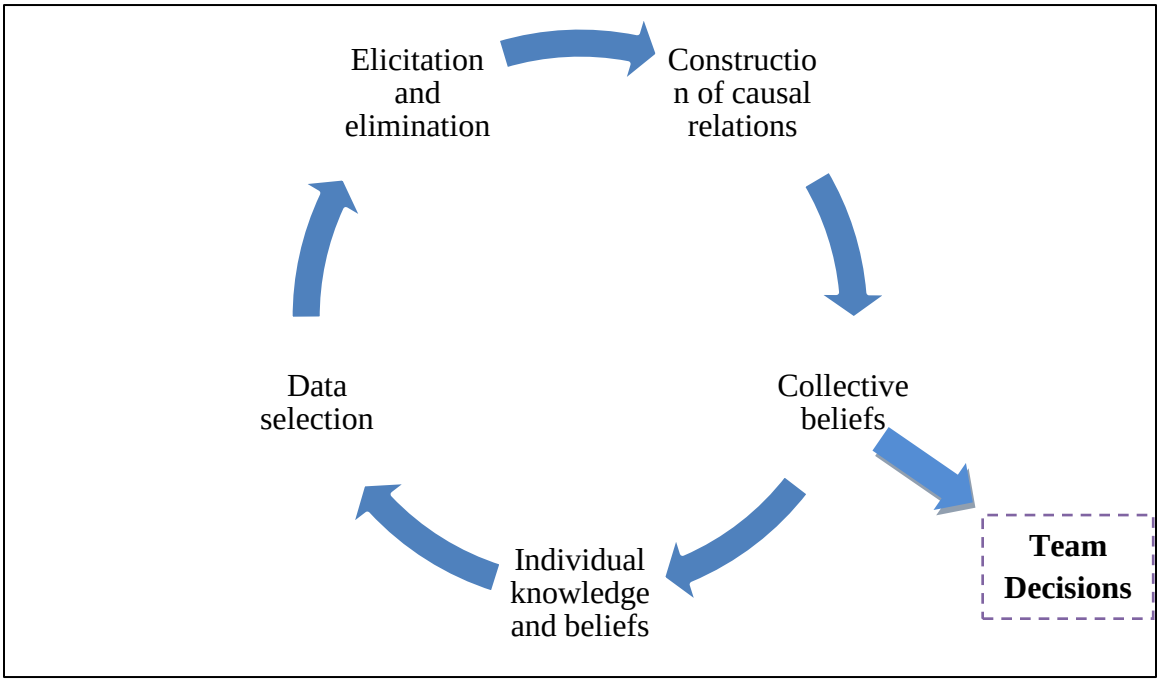

Figure 6 Collective common ground building (Deshpande et al. 2005)

Yi, Kang, Stasko, & Jacko (2008) raise three characteristics of sensemaking that are also relevant to this study. First sensemaking process is cyclic and iterative as the elaboration and reframing cycles imply. Second sensemaking is about creation as much as it is about exploration. The interactive conception of causal map is a collective creation process while individual frameworks that are results of elaboration and reframing are individual creations. Third sensemaking efforts are retrospective that is people do not create their frameworks upon information exploration but they initially create a framework and then look for the relevant information and verify whether the information confirms their expectations.

The discussion above highlights two major aspects of the sensemaking process. First there is a common ground which represents the understanding at a collective level. This serves as the basis on which the discussions are built implying the second, namely the longitudinal Common ground is not stable as all collective actions accumulate new information yielding to a continuous update of individual frameworks which in turn shifts the understanding at a community level from one state to another. Thus, in a simplistic way, sensemaking has a stationary and a continuous aspect.

II) CONTENT ANALYSIS AND CLUSTERING IN OBSERVING BASES

A discussion on a wicked problem is very likely to cover diverse topics as it is an intertwined network of problems from diverse fields. Each topic of discussion can therefore be considered as a base, each of which constitutes a different starting point for sensemaking. Content analysis is used to detect diverse topics of discussion. Content analysis entails the use of words that are present in discussions in order to distinguish documents with different focuses. The idea originates from document retrieval where a user query is matched against a text document. Internet search algorithms work with the same principle where the matches are sorted afterwards in terms of their relevancy using algorithms such as PageRank (L. Page, Brin, Motwani, & Winograd, 1999).

II.1) THE VECTOR SPACE MODEL

The Vector Space Modeling (VSM) (Salton et al., 1975) provides the guidelines for content analysis part of this study. Its major proposition is that any text document can be represented as a vector with dimensions corresponding to certain terms where T_i corresponds to the frequency of a term within a document (see **Figure 3**).

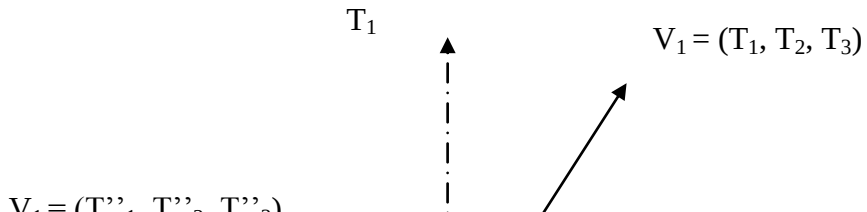

Figure 7 Vector representation of a document space

While a document space as in **Figure 4** would be the most desirable configuration, a practical way of assessing whether a document is relevant requires a priori information on the dimensions. Therefore an alternative that the VSM suggests is to obtain maximum separation among all documents as in **Figure 5**.

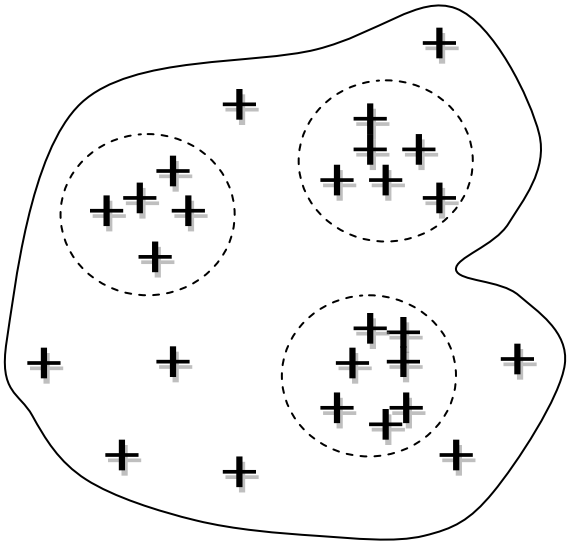

Figure 8 Desired document space configuration

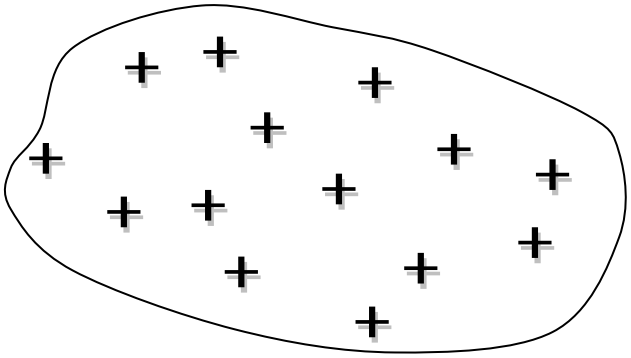

Figure 9 Document space with maximum separation of individual documents

In order to see which set of terms constitute dimensions yielding to a document space as in **Figure 5** the terms with the highest contribution in decreasing the density are selected where the document space density is defined as:

$$Q = \sum_{i=0}^{n} s(C^*, V_i)$$

where Q stands for the sum similarities (i.e. "s") between each document V_i and the centroid of the whole document space C^* defined as follows:

$$C^* = 1/n \sum_{i=0}^{n} V_i$$

The similarity between two vectors is calculated in function of the angle between them. The division of the inner product of two vectors with their magnitudes gives the cosine value of the angle in a range from 0 to 1. As the angle approaches 0, the cosine value, i.e. the degree of similarity moves toward 1.

Then the contribution of each term can be calculated as:

$$K_j = Q - Q_j$$

where,

$$Q_j = \sum_{i=0}^{j-1} s(C_j, V_i) + \sum_{i=j+1}^{n} s(C_j, V_i) \text{ and,}$$

$$C_j = {}^{1}/_{(n-1)} \left[\sum_{i=0}^{j-1} V_i + \sum_{i=0}^{j+1} V_i\right]$$

Thus terms that have a higher K value are more contributive in spreading documents across the space, in other words, in their absence the space gets more compact (i.e. Q increases). Therefore terms with positive K values are contributive while those with negative values are not further taken into account in clustering the documents.

However for better results, the way term frequency is computed should be refined because the mere count of terms does not reflect the real representative power of the term. First in a longer document a term tends to exist more. A term also loses from its representativeness with an increasing number of documents in which it exists. The latter issue implies also the elimination of stop words, i.e. words that are used very commonly or words that have no real sense , e.g. the conjugations of basic verbs such as “to be” or connecting words such as “and”, “or” etc. In order to alleviate the influence of stop words they are excluded from consideration right at the beginning by omitting from the initial set of terms that constitutes the input for the VSM. In order to lessen the influence of document length and overall term frequency within the space, the term frequency, is normalized. Normalized term frequency, i.e. w is defined as:

$$w = tf * \log({}^{N}/_{n})$$

II.2) STEMMING

Stemming stands for reducing words to their stems. Stem is not necessarily identical to the morphological root of the word. For information retrieval purposes it is enough to reduce related words to the same stem. Given that someone querying “hunt” would also be interested in “hunter” or “hunting” stemming, stemming has been implemented in Internet search engines as

well (e.g.(Dembo, 2008)). Therefore it provides a means to consider together different forms of a word such as adjective (speedy), adverb (speedily), name (speed) etc. This enables better handling of the content given that it permits decreasing the number of dimensions to consider in calculations, increasing the computation speed. However stemming is not flawless and needs to be applied with caution. It may end up stemming two irrelevant words to the same root (i.e. overstemming) or might fail to stem to the same root even though it should (i.e. understeming).

In this study, Porter's Stemming Algorithm for Italian is used (Porter, 2001). The abovementioned problems hold also for the algorithm for the Italian language. For instance both "moderne" i.e. modern and "moda" i.e. fashion are stemmed to "mod" although they are completely different. While "andare" is stemmed to "andar", its past participle form "andato" is stemmed to "andat". Synonymous words can also have completely different stems although they have exactly the same meaning while homonymous words bear different meanings in different contexts although they have exactly the same stem.

II.3) CLUSTER ANALYSIS

Clustering provides with a means to revel the community structure in a network. Although networks are usually understood as groups of interacting people, terms can form networks as well permitting to observe the stationary aspect of sensemaking. The terms obtained through the stemming algorithm explained above are therefore considered as nodes and edges correspond to the terms coexisting within the same document, i.e. a participant's online post. Thus the cluster of terms will provide the bases, i.e. different topics of discussion, on which the online community will build understanding through separate sensemaking processes.

Although various algorithms are suggested (e.g. Girvan & Newman (2002); Krause et al. (2003)), the agglomerative algorithm of Clauset, Newman, & Moore (2004) is preferred for the sake of computational ease. Initially all nodes are considered as separate entities among which connections are formed in function of the modularity measure

$A_{vw} = \begin{cases} 1 \; id \; vertices \; v \; and \; w \; are \; connected \\ 0 \; otherwise \end{cases}$ where A_{vw} is a member of the adjacency matrix of the network.

Supposing the network is split into clusters such that vertex v is in c_v. Then the ratio of vertices within the cluster to the total number of vertices in the network is:

$\frac{\sum_{vw} A_{vw} \, \partial(c_v c_w)}{\sum_{vw} A_{vw}} = \frac{1}{2m} \sum_{vw} A_{vw} \, \partial(c_v c_w)$ where $\partial(c_v c_w)$ is equal to 1 if i=j and 0 otherwise and $m = \frac{1}{2} \sum_{vw} A_{vw}$ is the number of edges in the network. The modularity is obtained by subtracting the expected number of edges within a cluster when they are assigned at random across vertices.

The degree k_v of a vertex v is the number of edges connected to that vertex:

$k_v = \sum_w A_{vw}$ The probability of an edge to exist between v and w in a network of randomly assigned edges while respecting vertex degrees is $k_v k_w/2m$. Then the modularity is:

$Q = \frac{1}{2m} \sum_{vw} \left[A_{vw} - \frac{k_v k_w}{2m} \right] \partial(c_v c_w)$. Modularity implies that the way a network is divided into cluster makes sense if the fraction of edges falling in a cluster is greater than the same figure expected in a random network. Therefore at each step, changes in modularity (ΔQ) are computed for each possible connection between pairs of clusters. Choosing the largest of them, the process continues until a single cluster is obtained while the number of clusters is decided by selecting

the solution yielding the highest Q. Therefore at step 0 there are n clusters in a network of n vertices each one formed by single individuals.

The way the term network is created has the assumption that two terms that are present within the same post deal with the same issue. Although inappropriate words are discarded by eliminating stop words and the use of vector space modeling, copresence itself does not tell much about the level of relatedness of two terms. A pair of terms that coexist in one post and another in a hundred of them should not be considered the same. Therefore the edges that are formed should be associated with weights that are simply the number of documents in which two terms are present together. With a little modification of the modularity measure as suggested by Newman (2004) the information held by weights is taken into account. Therefore A_{vw} is the now defined as the weight of the edge connecting vertices v and w while the degree of a vertex v is still: $k_v = \sum_w A_{vw}$ whereas the degree is modified as the sum of weights of edges that are incident on vertex v. The clusters obtained by this way are assumed to be representative of the topics that the online community discussed.

To observe the longitudinal aspect of sensemaking there is need to categorize messages posted by individuals in function of the topic they touch upon. This aims at observing the influence of previously posted messages about a certain topic on future contributions that are made within the same domain. This however, requires associating user posts with one of the term clusters. This can be done following a simple logic:

$V_h = \{tf_1, tf_2,\ldots,tf_i,\ldots,tf_n\}$ where v_h is a user post expressed through the frequency of the set of n terms used also in clustering where every $tf_j \in \{C_1, C_2,\ldots,C_j,\ldots,C_t\}$ and t is the number of clusters obtained from the clustering analysis above. Therefore membership of terms to a

particular cluster k can be expressed through a 1 x n vector $M_k = \{X_{1k}, X_{2k},\ldots,X_{ik},\ldots,X_{nk}\}$ where

$$X_{ik} = \begin{cases} 1 \; if \; the \; term \; Xi \; is \; associated \; with \; cluster \; C_k \\ 0 \; otherwise \end{cases}$$

A user post can be assigned to the cluster that is represented more than the others. The representation of a cluster within a post can be measured by the total frequency of the terms belonging to that cluster which can be defined as follows:

$$R_{hj} = V_h . M_k$$

Thus every post h is assigned to cluster j which provides the highest R value.

II.4) LONGITUDINAL ANALYSIS

Once every post is assigned to one of the clusters it is possible to examine whether a user's contribution in a particular cluster has anything to do what he/she read in the past. Here the assumption is that in case a user makes readings on a particular topic, he/she will have a ground to build a contribution within the same field. This may work through various mechanisms. Knowledge acquired through former readings can enable one to make a synthesis of different pieces of information and make a contribution by integrating his/her personal information. However a contribution does not have to be based on an approval of opinions presented in the online discussion. Even when a user does not agree with what has been said by the community or by certain individuals within the community, he/she can make an opposing claim with appropriate evidence. Commonsense does not necessarily mean everyone being in perfect agreement but it means that one understands the same thing as the rest of the community from a knowledge shared or a message sent. Actually a discussion where opposing views challenge each other can be productive in exploring the topic of discussion unless it turns into a conflict. Thus

there should be a connection between the number of posts that a user reads and the number of posts he/she authors. Given that there may be different bases on which the common sense is built, longitudinal analysis is done by investigating this relationship for every cluster where a cluster stands for a topic of discussion.

As the longitudinal analysis method, Constant Coefficients Modeling is used. This model assumes that regression coefficients are the same for all cases in the sample at all times. Thus it is convenient for data covering a relatively short period of time. It allows analyzing data that do not conform to Ordinary Least Squares (OLS) assumptions where homoskedasticity and the absence of serial correlation. The former stands for the assumption that the variance of the random disturbances is a constant that is valid for each case in the model. The latter indicates that the error terms associated with the same individual but with different time periods should not be correlated, i.e. there should not be serial correlation.

Next chapter presents the experiment held at the University of Naples Federico II in December 2007. Using data from this experiment, the aforementioned cluster and longitudinal analyses are conducted.

CHAPTER IV EXPERIMENT AND RESULTS

I) EXPERIMENT DESIGN

This study examines the hypothesis that effective collective intelligence can be achieved even by large scale online communities aiming at solving a wicked problem. This is considered to be possible by means of large-scale argumentation systems. The online argumentation tool that is used in this study is called Deliberatorium. Like other argumentation systems, it integrates sharing and argumentation technologies to enable the systematic identification of solutions on a large scale, and then use funneling to help participants come to consensus about which of these solution ideas should be implemented. However creating such large-scale argumentations systems faces challenges such as:

- Editing: Editing has been done in various ways in different online collaborative media. While Wikis allow edit anything, Slahdot.com has classes of editors with different authorizations. The wiki model is useful in attracting many different perspectives. However it suffers from unknowledgeable users' contributions that decrease the credibility of the online encyclopedia thereby has the potential risk of discouraging experts from participating. The forum model on the other hand supports expert commentary. However the followers in the thread have much less effort in influencing the quality of content. The online argumentation platform used in this experiment i.e. Deliberatorium, lets only the author of a post edit his/her own post while other users can submit their opinions for further consideration by the author. Any user can rate a suggestion aiming at providing with evaluation and guidance for the authors. Given that every opinion is represented by a single post in argumentation enabling open authorship

is not necessary. This also helps to reduce flame wars by not constraining the debate on a single entry.

- Quality: Deliberatorium lets users submit their contributions through informal comments and more structured arguments. Users can find the relevant argument tree branches through a search tool while there is additional information provided on the level of activity of the branches by highlighting more active branches in larger font. A special class of users called editors can replace user contributions by attaching them to other relevant posts. Editors however are supposed to be neutral in face of discussions. This design choice can provide a meaningful argument map at the outset of discussions and help users to post their contributions at the right parts of the content. Editors are responsible to correct the mistakes of placements. The activity scores are also considered as facilitators in guiding the community to convergence. The same content can be submitted in different ways in an argument map. However users are believed to prefer the most active branch in order to render their opinion more visible to the community.
- Consensus: Argumentation presents alternative solution methods in different branches of a map where each one combines a set of ideas different from the others. Rendering alternatives more visible to the community is a prerequisite for attaining consensus.

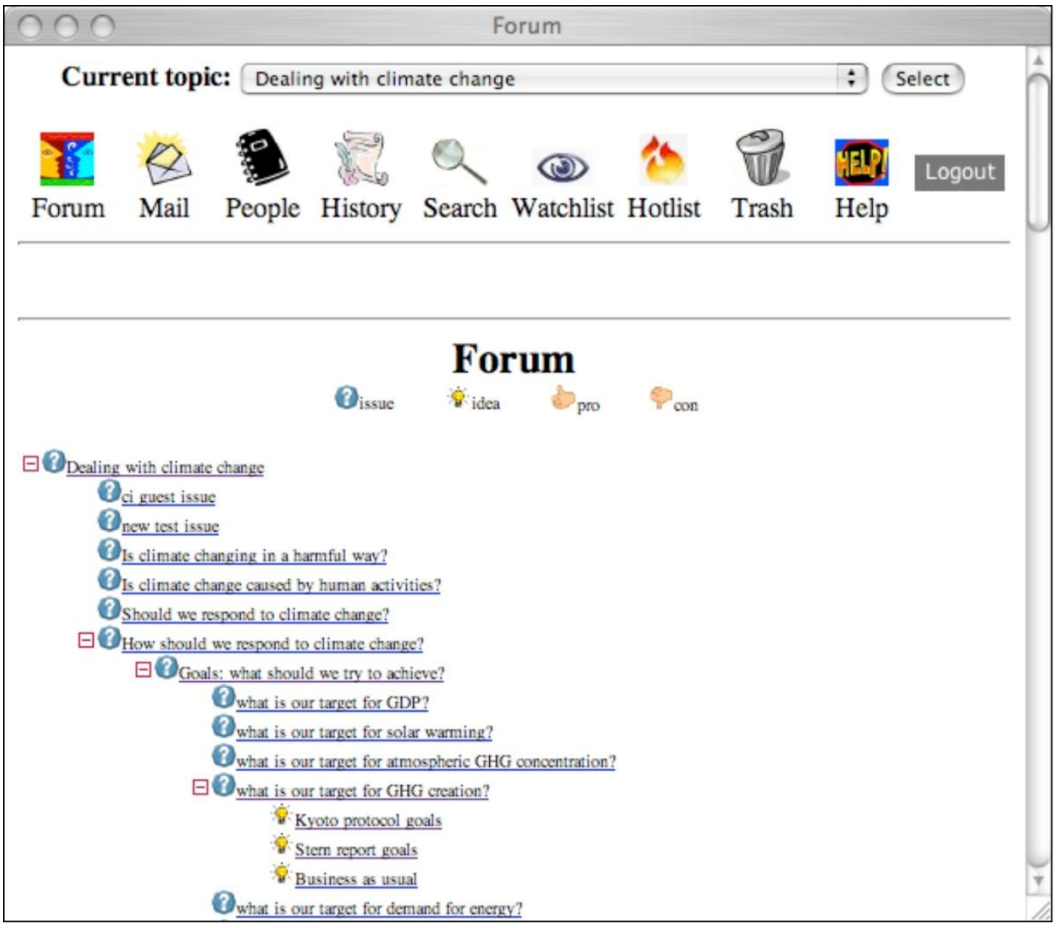

Figure 10 Snapshot from Deliberatorium

Several studies (e.g. (Buckingham Shum et al. 2004; De Moor & Aakhus, 2006) have targeted to effectively use the Internet technologies to implement argumentation in order to improve the quality of collective works in terms of knowledge sharing, representation and transfer. The intrinsic benefits of argumentation are used in order to:

- To represent knowledge in a compact way
- To let users relate a piece of information to the rest of the content by creating knowledge networks
- To improve the support for evidence thereby creating incentive for submitting more plausible posts

Deliberatorium integrates IBIS approach with Toulmin's argument analysis structure and Walton's schemes. The design aims at benefiting from the structure of argumentation as much as possible while keeping formalization at a minimum level. By merging the IBIS, Toulmin and Walton approaches, arguments are depicted as networks. An argument network is composed of nodes (claims) and edges (relationship between claims). A claim can be the premise or conclusion of an argument and can be considered true to a certain degree (e.g. based on the level of consensus assigned to it by an audience). An arc links two claims, specifically a premise to a conclusion and transfers the degree of truth of the premise to the conclusion. Arcs are associated with schemes through which they transfer the truth from the premise to the conclusion. The scheme assigned by users describes the way the conclusion is "inferred" from the premise.

I.1) SOFTWARE IMPLEMENTATION OVERVIEW

The Deliberatorium is a Common Lisp application developed on top of cl-http, an open source Web server developed at MIT (http:/www.cl-http.org:8001/). It provides a simple and consistent web-based user interface that allows users to navigate and edit the argument map as well as communicate with each other. The system's capabilities is made accessible via a set of tool icons arrayed across the toolbar at the top of the page (**Figure 3**). The tools include:

- The argument map: This allows users to browse and edit the argument map. The argument map aims at allowing users to get a sense of what other community members are doing, thereby fostering emergent self-organization. This is achieved by providing visual cues concerning which branches of the argument map are most active, which posts are the most highly-rated etc.. The edit history for all articles is kept in the system. Therefore one can quickly see the former editions of an article if desired.

- Search and history: These allow users to find the posts that have given keywords or edit histories or were looked at recently by that user
- People and Home: every user has a customizable home page which lists which articles and comments they have contributed. These enable developing an online presence if desired, facilitating reputation-building, networking, and community-building. The people tool provides links to the home pages for all the users registered for the current topic.
- Mail, Chat room and Forum: These tools allow one to one communication (mail), in a public synchronous context (chat room), or via a public asynchronous threaded discussion (forum).
- Watchlist: This allows users to specify which articles or comments they are interested in, so they can be automatically notified by email when any changes are made.
- Survey: These allow users to provide feedback on the system.
- The help tool: This provides a set of textual user guidelines as well as help videos.

I.2) THE SET UP OF THE EXPERIMENT

A first test of the Deliberatorium was performed in December of 2007 at the University of Naples Federico II (Italy) with a community of more than 150 users. Although the number of user accounts created was more than 200 some of them were duplicate accounts. This figure finally decreased to 165 users including editors after the verification with participants. Users were graduate students who were asked to deliberate on the topic "the future of biofuels". The students were all part of a same class from a graduate program in Industrial Engineering, age 23-25, 55% male. Students selected from that class helped to coordinate and manage the experiment, and were required as a result to deal with social pressures from their fellow students

and with the fact that most students inevitably felt the experiment was a course task for which they will be evaluated by their professor. All these circumstances made the context different from a fully open online community and represent a significant limitation of this study. However going large scale in the early steps within an uncontrolled experimental setting was undesired in face of the risk of not attracting a critical mass of users.

The test developed in four phases, starting from early November 2007:

1. Preparatory work
2. A three weeks period, in which students were requested to populate the Deliberatorium with contents
3. One week for consolidating the knowledge map produced by the community
4. Data analysis

In the preparatory phase, the students had four 2 hours seminars from external experts about:

- collective intelligence and its current internet applications
- argumentation, with focus on the IBIS approach
- major issues in energy governance with a country focus on Italy and UE policies
- an instructional demo of the Deliberatorium beta version. The students were also given a few reading materials: two newspaper and magazine articles about the topic and the IBIS manual available at http://touchstone.com/wp/IBIS.html (J Conklin, 2009)

The criteria used to select the "future of biofuels" as a topic were: First the topic had to present a wicked problem such as global warming. Second it had to be focused enough to help students not get lost into a too wide a debate, given that they had limited time, attention and expertise. Finally

it had to be controversial so that the community could explore possible different solutions and perspectives.

Instead of giving students an empty argument map, the following first level questions and options were provided at the beginning: 1) what percentage of transportation energy needs in Italy will come from biofuel consumption twenty years from now? (Options: limited (less than 20%), moderate (between 20 and 30%), substantial (more than 30%)); 2) how can Italy get the biofuels it needs? (No option provided).

I.3) THE COMMUNITY

In the Deliberatorium community, there are three roles: moderators, authors and readers/voters. Moderators are charged with the usual tasks of filtering out noise and rejecting off-topic posts. They were also in charge of ensuring that the argument map was well-structured, i.e. that all posts were properly divided into individual and non-redundant issues, ideas, and arguments, and were located in the relevant branch of the argument map. This involved classifying and sometimes editing posts, offering suggestions to authors, aggregating similar arguments, and occasionally re-organizing the overall argument map so that related topics are grouped into the same branch. A team of 4 student moderators was selected and trained in argument mapping before the test. The on-line argumentation process developed as follows:

- Authors posted and edited questions, ideas, and pro/con arguments and produced an argument map similar to that in **Figure 2**. While questions and ideas could be posted only as single short sentences, arguments were posted using an on-line form that helped them structure their post in argument form (conclusion, argument scheme and critical questions, argument content, possibility to attach links, references and documents); the form was

designed to facilitate users' contribution while preventing them from submitting unstructured content.

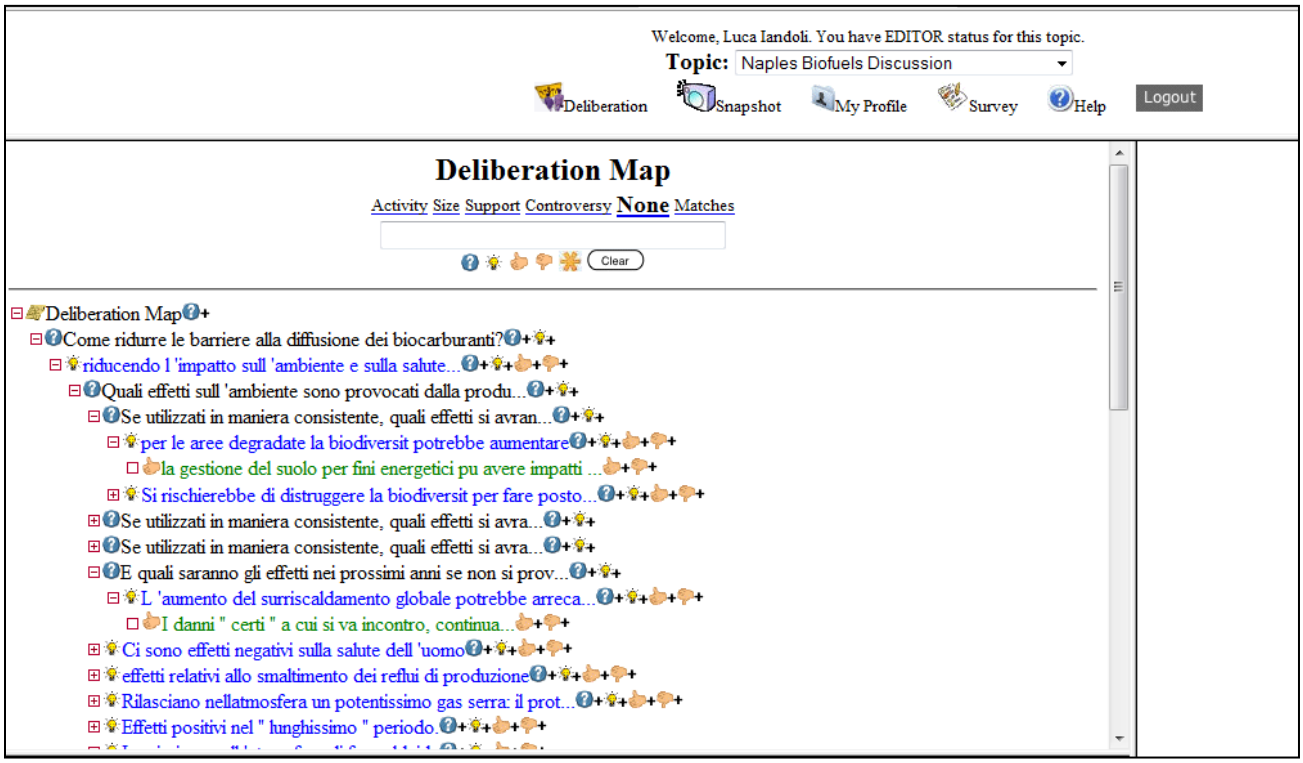

Figure 11 An example of argument map from Deliberatorium

- All users (including moderators, authors and readers) rated arguments and ideas and could send comments to authors through threaded discussion forums associated, like wiki talk pages, with each post. Rating was anonymous;
- Posts were initially given a status of "pending", and could only be certified by moderators. Until a post was certified, it could not be rated and nobody, except its author, could link any other posts to it. We also explained that only certified posts will appear in the final, publicly available, version of the argument map. Moderators also left comments, edited, moved, trashed and classified posts. Usually moderators would leave a comment to explain changes. Authors would receive an alert email when their post was modified or trashed (but the trash

was never emptied). In the experiment nobody, except moderators, was allowed to edit a post authored by someone else.

- Several countermeasures and incentives were set up to limit the negative effects due to limited scale and presence of social and informational pressure usually absent or limited in Internet communities. In particular, with the support of Naples City Science Museum extrinsic incentives such as minor awards were provided to improve the post quality. To limit the negative influence of social pressure on the rating process, a kind of prediction market incentive for voters was set up according to which votes would have been converted at the end of the experiments into awards. For this purpose at the end of phase 2, a team of independent experts would have identified and ranked the best posts. Then voters would have been assigned a score based on how closely their votes correlated with the expert ratings. The voters with the highest correlation score would have been selected and awarded with educational gadgets.

II) RESULTS

Deliberatorium was active almost 24 hours per day, except for a hiatus between roughly 3 and 6 am. In two weeks the community posted nearly 3000 issues ideas and arguments (of which roughly 1900 were eventually certified) in addition to over 2000 comments. They were, however, relatively few ratings, notwithstanding the presence of extrinsic incentives: each post received an average of only 2.2 ratings. The intensity of participation reminds the Power law distribution (**Figure 3)** that has been found to be typical of many on-line communities below (Madey, Freeh, & Tynan, 2002).

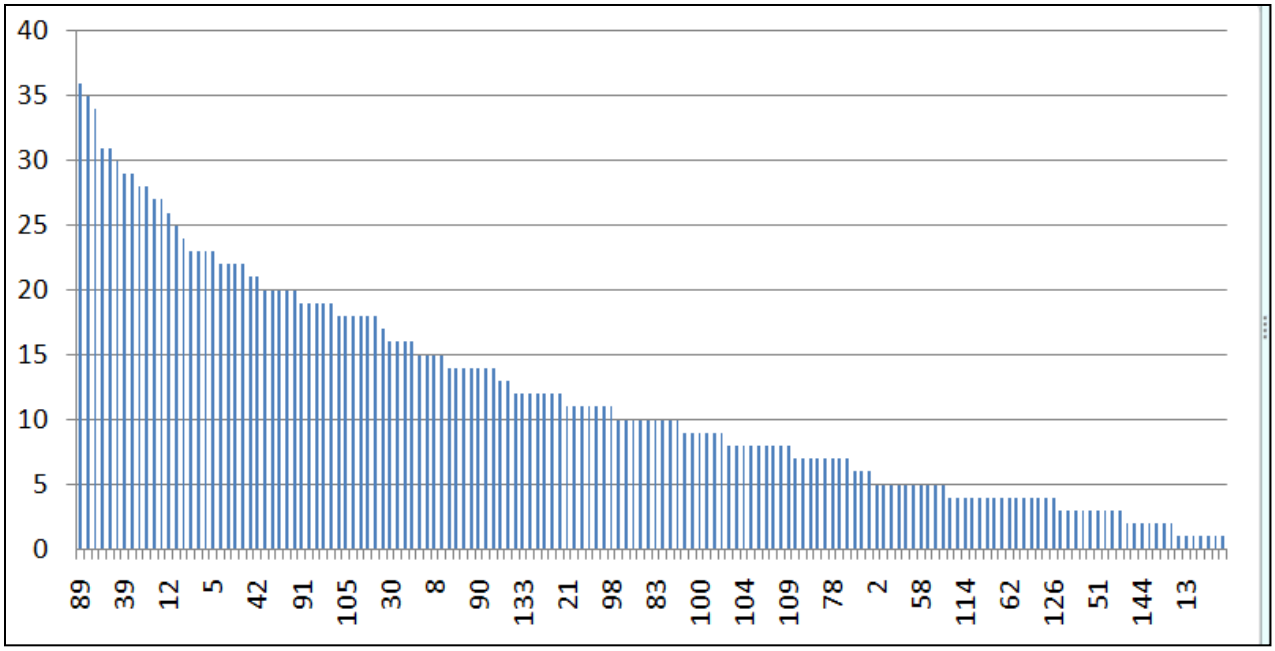

Figure 12 Number of posts authored by users

The breadth of coverage was satisfactory as a non-expert community of students was able to create a comprehensive map of the current debate on biofuels in just a couple of weeks, exploring topics ranging from technological issues to environmental, economic and socio-political impacts of the widespread diffusion of biofuels. Moreover, the proportion of out-of-topic posts was negligible – about 0.1%.

The dominant argument scheme was "by authority", followed by analogy, deductive, and inductive schemes. It also appears that users were generally not able to associate the right scheme to their arguments, which increased the moderators' workload. Though students participation may have been influenced by their perception that the experiment was a course task for which they could be evaluated by their professor, their informal face-to-face and on-line comments, posted on the Deliberatorium as well as on a threaded discussion forum run

independently by a students' association web site, showed that they found the experiment interesting and appreciated the innovative characteristics of the Deliberatorium.

II.1) SENSEMAKING OBSERVATIONS

As mentioned also in Chapter III, sensemaking has two dimensions, one stationary, another longitudinal that are also named as basis and accumulativeness here. In solving a wicked problem it is expected to encounter issues related to different areas of expertise which implies that there should be different bases on which the discussions take place simultaneously. Therefore the presence of different bases is taken as a support for the existence of commonsense although it is worth noting that it does not guarantee achievement without accumulativeness. The content analysis conducted to observe bases, each one constituting a topic of discussion, consists of various steps.

1. Initially the set of messages to taken into account are selected. First, although there were various posts that were still waiting to be certified even at the end of the experiment, they were visible to the community. Therefore they had influence in transmitting knowledge, so they are also taken into account. Second, some posts were submitted by nusers more than once. This might have been due to the author's confusion or to the fact that the author can resubmit a rejected post without modifying it, maybe hoping that it would be edited by a different moderator or he/she can resubmit by making minor modifications. Given that they will fall into the same topic of discussion they will have identical effects as well. Therefore they are retained in the analysis. Third the posts authored by editors are not taken into account in order to observe the user group that resembles most the situation in an open online discussion community. Finally, not being a part of the

argumentation, informal content that was posted in the chat room, the forum or in the comments were discarded. In the end there were 1839 posts that are taken into account.

2. The set of posts selected provided with a large set of words that is used to construct a network of terms where two terms are connected with an edge when they are used within the same post. However to get more meaningful results, all words are stemmed using the Porter's (2001) stemming algorithm for the Italian language. It is written using Python scripting language which yielded a set of 6783 stems where different forms of a word such as adjective, adverb etc. are considered as a single entity.
3. The set of stems constituted the set of terms replacing the initial set of words. However as explained by the Vector Space Model (Salton et al., 1975) some words are not that helpful in obtaining the clusters representing the topics of discussion. Therefore the set of terms is refined by using this algorithm still using the Python language.
4. The refined set of terms is then used to construct the network where nodes represent terms. Two terms are connected with an edge if they appeared within the same post. However the level of relation between pairs of terms cannot be taken identical given that some may be present together in several posts whereas some may appear just once. Therefore the network edges are weighed by the number of posts in which the connected terms are present together.
5. Following the construction of the network, the clustering algorithm of Clauset, Newman, & Moore (2004) is used which yielded four major stem clusters (see **Table 1**).

Table 5 Stem Clusters

Cluster No	Number of Stems
1	1788

2	2279
3	2353
4	286
5	7
6	3
7	9
8	13
9	7
10	7
11	5
12	12
13	6
14	8
Grand Total	**6783**

As Hair et al. (1995) mention, clustering is the most subjective statistical technique. Therefore it is not surprising to obtain meaningless clusters. The low number of stems in many clusters point out this fact. Thus only the top 3 clusters are taken into account initially where each one has somehow different contents (see **Table 2** in Introduction). For instance cluster 1 is composed of economics and political terms while cluster 2 has political, organizational terms. Finally cluster 3 has environment and technology related stems. For the sake of prudence cluster number 4 is also taken into further consideration in the assignment of user posts to every cluster although the stems in this cluster did not constitute a meaningful set.

6. Once the cluster s of stems are obtained, it is possible to assess the relevance of every user post with respect to each cluster, representing a different topic. Every post is associated with the cluster that is represented the most i.e. the total number of occurrences of the terms associated with that cluster is higher than the other clusters' terms. Among four clusters taken into account three of them are retained for further analysis given that the fourth cluster of stems was associated with only 22 user posts

whereas the next least represented cluster was represented through 314 posts (see **Table 2** below).

Table 6 Cluster - User post assignments

Cluster No	Number of user posts associated with the cluster
0	314
1	563
2	940
3	22

As the cluster analysis suggests, there are diverse topics of discussion, majorly three of them therefore there is reason to believe that sensemaking process can be built on these grounds. In order to analyze the longitudinal aspect of the sensemaking process, Constant Coefficient Model is used according to the following model:

$$\mathrm{cr}_{ijt} = \beta_{0i} + \beta_{1i} \log(vw_{ijt}) + \sum_{j=0}^{j=N} \beta_{ji} D_{jt} + u_{jt}$$

Where i €{1,2,3} the set of clusters; j €{1,…,N} the set of users, t=1,2,3 the periods that represent the first, second and third weeks of the experiment respectively. Then the variables are:

$$D_{jt} \begin{cases} 1 \; if \; user \; j \; at \; time \; t \\ 0 \; otherwise \end{cases}$$

cr_{it} = number of posts authored by user j at time t on a topic represented by cluster i

vw_{ijt} = number of times user j viewed another user's post on topic i at time t

Note that this model is run for every cluster separately by taking into account the serial correlation which assumes that the errors at time t are correlated with the errors in the previous

time period t-1 where the serial autocorrelation parameter is equal for all individuals j in the sample. Therefore: $u_{jt} = \partial u_{it-1} + €_{it}$ where the parameter ∂ is constant across individuals and over time while € is the white noise.

The model is tested using Stata. The results yield affirmative results on the longitudinal aspect of sensemaking where previously read posts are found to have influence on the contributions of the user on that same topic.

Table 7 Constant Coefficients Model with serial correlation tested for each cluster

	Variable	Coefficient	Std Error	P>|z|
Cluster 1	vw	.665	.031	0.00
	constant	.345	.018	0.00
Wald chi2(1) = 459.99, Prob > chi2 = 0.00, Number of observations = 495, Number of groups (i.e. users) = 165, Time periods = 3				
	Variable	**Coefficient**	**Std Error**	**P>|z|**
Cluster 2	vw	1.226	.021	0.00
	Constant	.483	.019	0.00
Wald chi2(1) = 3521.32, Prob > chi2 = 0.00, Number of observations = 495, Number of groups (i.e. users) = 165, Time periods = 3				
	Variable	**Coefficient**	**Std Error**	**P>|z|**
Cluster 3	vw	1.969	.039	0.00

Constant	.292	.029	0.00
Wald chi2(1) = 6.87, Prob > chi2 = 0.0088, Number of observations = 495, Number of groups (i.e. users) = 165, Time periods = 3			

II.2) DISCUSSION

This study uses sensemaking approach in order to assess argumentation's capability in mediating online wicked problem solving. Such efforts are challenging given the presence of diverse fields to deal with as well as the need to gather people from those areas of expertise. This study used data from an experiment which can be considered as an initial attempt to gather large number of people with the objective of solving a wicked problem. Due to the significant lack of use of argumentation in such context, a student community was considered useful in order to observe the interaction of users and the effectiveness of argumentation. Although the community was not composed of experts in biofuels or any other related field, the level of participation gave encouraging results for the use of argumentation in a real world situation.

The so called stationary dimension of sensemaking is shown to exist with the presence of diverse discussion topics. Therefore different points of departure are provided for building a collective understanding. As sensemaking concept suggests, it is a process rather than an event taking place at a certain point of time. Therefore the process of sensemaking is examined for every cluster through time. It is found that a user's past readings on a certain field leads him/her to make contributions on the same field.

Although there is support for the effectiveness of argumentation on enabling sensemaking, sensemaking alone is not enough to come up with a successful solution to a wicked problem. There are other requirements such as the presence of knowledgeable people, resources such as

time, money etc. However in the absence of the ability to build a common sense the abundance of these resources would not be enough either. Therefore there is need for further research to improve the conditions for attaining an acceptable wicked problem solution.

The approach used in this study can also constitute a point of departure for the comparison of alternative platforms with varying designs. However one should still keep in mind that a tool's ability to support sensemaking better than others does not make it the best option in the presence of other requirements mentioned above.

II.3) **LIMITATIONS**

The study comes along with its limitations as well. In addition to the aforementioned limitations on the experiment design, there are some methodological concerns that are worth mentioning. First, although stemming facilitates regrouping related words in a single group, in the end it is an algorithm and it is not 100% reliable. As there is no language with Cartesian grammar rules, there is no perfect algorithm to reveal the morphological structure of every word.

Second clustering is a subjective approach in revealing the structure of a network although it is widely used in network analyses. Different clustering methods yield different results. The major reason for selecting the clustering algorithm of Clauset et al. (2004) has been the computational limitations. Other algorithms such as the one suggested by Newman & Girvan (2003) may take more than 20 hours with the dataset used in this research. Clustering may provide groups with limited number of entities that can be discarded or can be united with another group. The imperfect results of stemming also has consequences on clustering where not so related stems that are outputs of stemming become inputs for clustering. Although stems and clusters are carefully looked through, people may come up with different decisions.

BIBLIOGRAPHY

Ackoff, R. L. (1973). Science in the Systems Age: Beyond IE, OR, and MS. *Operations Research*, *21*(3), 661-671.

Ackoff, R. L. (1956). The Development of Operations Research as a Science. *Operations Research*, *4*(3), 265-295.

Bailey, D. E., & Barley, S. R. (2005). Return to work: Toward post-industrial engineering. *IIE Transactions*, *37*(8), 737–752.

Barnett, W. P. (2008). *The red queen among organizations*. Princeton University Press.

Barron, B. (2003). When smart groups fail. *The journal of the learning sciences*, *12*(3), 307.

Bar-Yam, Y. (2002). Complexity rising: From human beings to human civilization, a complexity profile. *Encyclopedia of Life Support Systems (EOLSS), UNESCO, EOLSS Publishers, Oxford, UK*.

Beers, P., Boshuizen, H., Kirschner, P., & Gijselaers, W. (2005). Computer support for knowledge construction in collaborative learning environments. *Computers in Human Behavior*, *21*(4), 623-643. doi: 10.1016/j.chb.2004.10.036

Bonaccorsi, A., & Rossi, C. (2003). Why open source software can succeed. *Research Policy*, *32*(7), 1243–1258.

Buckingham Shum, S., Uren, V., Li, G., Sereno, B., & Mancini, C. (2004). Modelling Naturalistic Argumentation in Research Literatures: Representation and Interaction Design Issues. *International Journal of Intelligent Systems*, *22*(1), 17-47.

Campbell, V. (2004). How RAND Invented the Postwar World: Satellites, Systems Analysis, Computing, the Internet–Almost All the Defining Features of the Information Age Were Shaped in Part at the RAND Corporation. *Invention and Technology*, (Summer), 50-59.

Carlsson, B., & Jacobsson, A. (2005). On Contamination in Information Ecosystems. In *Hawaii International Conference on System Sciences* (Vol. 7, pag. 185b). Los Alamitos, CA, USA: IEEE Computer Society. doi: http://doi.ieeecomputersociety.org/10.1109/HICSS.2005.463

Checkland, P. (1985). From Optimizing to Learning: A Development of Systems Thinking for the 1990s. *The Journal of the Operational Research Society, 36*(9), 757-767.

Chidambaram, L. (1996). Relational Development in Computer-Supported Groups. *MIS Quarterly, 20*(2), 143-165.

Chidambaram, L., & Jones, B. (1993). Impact of Communication Medium and Computer Support on Group Perceptions and Performance: A Comparison of Face-to-Face and Dispersed Meetings. *MIS Quarterly, 17*(4), 465-491.

Clark, H. H., & Schaefer, E. F. (1989). Contributing to discourse. *Cognitive Science: A Multidisciplinary Journal, 13*(2), 259–294.

Clark, H. H. (1996). *Using language*. Cambridge University Press.

Clauset, A., Newman, M. E. J., & Moore, C. (2004). Finding community structure in very large networks. *Physical Review E, 70*(6), 066111. doi: 10.1103/PhysRevE.70.066111

Cohen, W. M., & Levinthal, D. A. (1990). Absorptive capacity: a new perspective on learning and innovation. *Administrative science quarterly, 35*(1).

Conklin, J. (2009). White Papers. *CogNexus Institute*. Recuperato Novembre 29, 2009, da http://cognexus.org/id26.htm#the_ibis_manual

Conklin, J. (2005). *Dialogue Mapping: Building Shared Understanding of Wicked Problems*. Wiley.

Cowgill, B. (2005 21). Putting crowd wisdom to work. *The Official Google Blog*. Recuperato da

http://googleblog.blogspot.com/2005/09/putting-crowd-wisdom-to-work.html

Cry Definition | Definition of Cry at Dictionary.com. (S.d.). . Recuperato Ottobre 13, 2009, da http://dictionary.reference.com/browse/cry?r=1

De Moor, A., & Aakhus, M. (2006). Argumentation support: From technologies to tools. *Communications of the ACM*, *49*(3), 93–98.

Delbecq, A. L., & Van de Ven, A. H. (1971). A Group Process Model for Problem Identification and Program Planning. *Journal of Applied Behavioral Science*, *7*(4), 466-492. doi: 10.1177/002188637100700404

Dembo, D. (2008, Novembre 21). Official Google Blog: Our international approach to search. Recuperato Novembre 23, 2009, da http://googleblog.blogspot.com/2008/11/our-international-approach-to-search.html

Dennis, A. R., George, J. F., Jessup, L. M., Nunamaker, J. F., & Vogel, D. R. (1988). Information Technology to Support Electronic Meetings. *MIS Quarterly*, *12*(4), 591-624.

Deshpande, N., De Vries, B., & Van Leeuwen, J. P. (2005). Building and supporting shared understanding in collaborative problem-solving. In *Information Visualisation, 2005. Proceedings. Ninth International Conference on* (pagg. 737–742).

Drucker, P. F. (2000). Knowledge-worker productivity: The biggest challenge. *The knowledge management yearbook 2000–2001*.

Fussell, S. R., Kraut, R. E., & Siegel, J. (2000). Coordination of communication: effects of shared visual context on collaborative work. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work* (pagg. 21-30). Philadelphia, Pennsylvania, United States: ACM. doi: 10.1145/358916.358947

George, J. M. (1990). Personality, affect, and behavior in groups. *Journal of Applied Psychology*,

75(2), 107-116.

Girvan, M., & Newman, M. E. J. (2002). Community structure in social and biological networks. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(12), 7821-7826. doi: 10.1073/pnas.122653799

Grossman, L. (2009, Giugno 17). Iran Protests: Twitter, the Medium of the Movement. *Time*. Recuperato Novembre 10, 2009, da http://www.time.com/time/world/article/0,8599,1905125,00.html

Heylighen, F. (1999). Collective Intelligence and its Implementation on the Web: Algorithms to Develop a Collective Mental Map. *Computational & Mathematical Organization Theory*, *5*(3), 253-280. doi: 10.1023/A:1009690407292

von Hippel, E., & Katz, R. (2002). Shifting Innovation to Users via Toolkits. *Management Science*, *48*(7), 821-833.

Holmes, J. G., & Rempel, J. K. (1989). Trust in close relationships. *Close relationships*, *10*, 187–220.

How to slake a planet's thirst. (2003). *Nature*, *422*(6929), 243. doi: 10.1038/422243a

Huber, G. P. (1984). Issues in the Design of Group Decision Support Sytems. *MIS Quarterly*, *8*(3), 195-204.

Hurwitz, R., & Mallery, J. C. (1995). The Open Meeting: A Web-based system for conferencing and collaboration. In *Proceedings of the Fourth International Conference on The World-Wide Web*.

J\osang, A., Ismail, R., & Boyd, C. (2007). A survey of trust and reputation systems for online service provision. *Decision Support Systems*, *43*(2), 618–644.

Janis, I. L. (1982). *Groupthink*. Houghton Mifflin.

Joseph F. Hair, J., Anderson, R. E., Tatham, R. L., & Black, W. C. (1995). *Multivariate data analysis (4th ed.): with readings*. Prentice-Hall, Inc.

Kesten Green, J. Scott Armstrong, & Andreas Graefe. (2007). Methods to Elicit Forecasts from Groups: Delphi and Prediction Markets Compared. *Foresight: The International Journal of Applied Forecasting*, Foresight: The International Journal of Applied Forecasting, (8), 17-20.

King, R. (2006, Luglio 8). Hollywood Games People Play. *BusinessWeek: Technology*. Recuperato Novembre 10, 2009, da http://www.businessweek.com/technology/content/aug2006/tc20060804_618481.htm?chan=top+news_top+news

Klein, G., Moon, B., & Hoffman, R. R. (2006). Making sense of sensemaking 2: A macrocognitive model. *IEEE Intelligent Systems*, *21*(5), 88.

Klein, G., Moon, B., & Hoffman, R. (2006). Making Sense of Sensemaking 1: Alternative Perspectives. *Intelligent Systems, IEEE*, *21*(4), 70-73. doi: 10.1109/MIS.2006.75

Klein, M., Cioffi, M., & Malone, T. (2007). Achieving Collective Intelligence via Large Scale Argumentation. *MIT Center for Collective Intelligence, Working Paper, July*.

Krause, A. E., Frank, K. A., Mason, D. M., Ulanowicz, R. E., & Taylor, W. W. (2003). Compartments revealed in food-web structure. *Nature*, *426*(6964), 282-285. doi: 10.1038/nature02115

von Krogh, G., & von Hippel, E. (2006). The promise of research on open source software. *Management Science*, *52*(7), 975.

Lindzen, R. (2006, Aprile 12). Climate of Fear. *Wall Street Journal*. Recuperato da http://www.opinionjournal.com/extra/?id=110008220

Linstone, H. A., & Turoff, M. (S.d.). *The Delphi method: Techniques and applications*.

Ludwig Von Bertalanffy. (1972). The History and Status of General Systems Theory. *The Academy of Management Journal*, *15*(4), 407-426.

Madey, G., Freeh, V., & Tynan, R. (2002). The open source software development phenomenon: An analysis based on social network theory. Presentato al Americas Confer- ence on Information Systems (AMCIS-2002).

Manyika, J., Sprague, K., & Yee, L. (2009, Ottobre 27). McKinsey: What Matters: Using technology to improve workforce collaboration. *What Matters*. Online Magazine, . Recuperato Novembre 4, 2009, da http://whatmatters.mckinseydigital.com/internet/using-technology-to-improve-workforce-collaboration

Markus, M. L. (1987). Toward a'critical mass' theory of interactive media. *Communication Research*, *14*(5), 491–511.

Marwell, G., & Ames, R. E. (1981). Economists free ride, does anyone else? : : Experiments on the provision of public goods, IV. *Journal of Public Economics*, *15*(3), 295-310. doi: 10.1016/0047-2727(81)90013-X

McAllister, D. J. (1995). Affect- and Cognition-Based Trust as Foundations for Interpersonal Cooperation in Organizations. *The Academy of Management Journal*, *38*(1), 24-59.

Newman, M. E. J., & Girvan, M. (2003). Finding and evaluating community structure in networks. *cond-mat/0308217*. Recuperato Marzo 6, 2009, da http://arxiv.org/abs/cond-mat/0308217

Newman, M. E. J. (2004). Analysis of weighted networks. *Physical Review E*, *70*(5), 056131. doi: 10.1103/PhysRevE.70.056131

Page, L., Brin, S., Motwani, R., & Winograd, T. (1999). The PageRank Citation Ranking:

Bringing Order to the Web. *STANFORD INFOLAB*, 17.

Page, S. E. (2008). *The difference: How the power of diversity creates better groups, firms, schools, and societies*. Princeton Univ Pr.

Pala, C. (2005). ENVIRONMENTAL RESTORATION: To Save a Vanishing Sea. *Science*, *307*(5712), 1032-1034. doi: 10.1126/science.307.5712.1032

Pelled, L. H., Eisenhardt, K. M., & Xin, K. R. (1999). Exploring the black box: An analysis of work group diversity, conflict, and performance. *Administrative Science Quarterly*, *44*(1), 1–3.

Porter, M. (2006). An algorithm for suffix stripping. *Program: electronic library and information systems*, *40*(3), 211 - 218. doi: 10.1108/00330330610681286

Porter, M. (2001). Snowball: A language for stemming algorithms. Recuperato Ottobre 13, 2009, da http://snowball.tartarus.org/texts/introduction.html

Rafaeli, S., & Raban, D. R. (2005). Information sharing online: a research challenge. *International Journal of Knowledge and Learning*, *1*(1/2), 62 - 79. doi: 10.1504/IJKL.2005.006251

Raymond, E. S. (2001). *The Cathedral and the Bazaar: Musings on Linux and Open Source by an Accidental Revolutionary*. O'Reilly \& Associates, Inc. Recuperato Novembre 3, 2009, da http://portal.acm.org/citation.cfm?id=560911

Rittel, H. W. J., & Weber, M. M. (1973). Dilemmas in a general theory of planning. *Policy Sciences*, *4*(2), 155.

Rosenhead, J., & Mingers, J. (2001). A new paradigm of analysis. In *Rational analysis for a problematic world revisited: problem structuring methods for complexity, uncertainty and conflict* (2° ed., Vol. 366, pagg. 1-19). Chichester (NY): Wiley.

Salton, G., Wong, A., & Yang, C. S. (1975). A vector space model for automatic indexing. *Commun. ACM*, *18*(11), 613-620. doi: 10.1145/361219.361220

Sperber, D., & Mercier, H. (2009). Reasoning as a social competence. In *Collective Wisdom*. MIT Press.

Sunstein, C. R. (2006). *Infotopia*. Oxford University Press US.

Taylor, F. W. (1914). *The principles of scientific management*. Adamant Media Corporation.

Toulmin, S. E. (2003). *The uses of argument*. Cambridge University Press.

Toulmin, S. E., Rieke, R. D., & Janik, A. (1979). *An introduction to reasoning*. Macmillan.

Tsoukas, H. (1996). The Firm as a Distributed Knowledge System: A Constructionist Approach. *Strategic Management Journal*, *17*, 11-25.

Van Valen, L. (1973). A new evolutionary law. *Evolutionary Theory*, *1*(1), 1–30.

Vennix, J. (1996). *Group Model Building: Facilitating Team Learning Using System Dynamics. ISBN 047195355*. Wiley, New York, New York.

Viégas, F. B., Wattenberg, M., & Dave, K. (2004). Studying cooperation and conflict between authors with history flow visualizations. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pagg. 575-582). Vienna, Austria: ACM. doi: 10.1145/985692.985765

Walton, D. N. (2006). *Fundamentals of critical argumentation*. Cambridge University Press.

Walton, D. N. (1989). *Informal logic*. Cambridge University Press.

Wilson, B. (2001). *Soft systems methodology*. Wiley.

Yi, J. S., Kang, Y., Stasko, J. T., & Jacko, J. A. (2008). Understanding and characterizing insights: how do people gain insights using information visualization? In *Proceedings of the 2008 conference on BEyond time and errors: novel evaLuation methods for*

Information Visualization (pagg. 1-6). Florence, Italy: ACM. doi:

10.1145/1377966.1377971