Re-engineering Knowledge Networks for Development

by

AMRITH RAGHAVAN

B.Eng., Industrial Engineering, University of Roorkee, India 1997
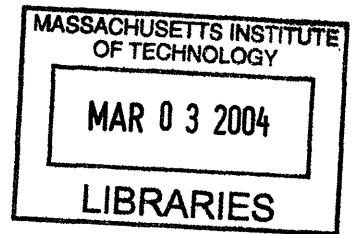MBA, Indian Institute of Management, Calcutta, India, 1999

Submitted to the Engineering Systems Division
in Partial Fulfillment of the Requirements for the Degree of
Master of Science in Technology and Policy

at the

Massachusetts Institute of Technology

February 2004

Signature of Author
Technology and Policy Program, Engineering Systems Division
January 26, 2004

Certified by
Nazli Choucri
Professor, Political Science & Associate Director, Technology & Development Program
Thesis Supervisor

Accepted by
Dava J. Newman
Associate Professor of Aeronautics and Astronautics and Engineering Systems
Director, Technology and Policy Program

Re-engineering Knowledge Networks for Development
by
Amrith Raghavan

ABSTRACT

This thesis examines the evolution of Internet based knowledge networks (e-knowledge networks) in the domain of sustainable development. The objective of this thesis is to use an engineering systems approach to understand knowledge networks, identify the barriers to their sustainability and recommend strategies for re-engineering them.

e-Knowledge Networks refer to the set of Internet based tools and platforms that support communication, collaboration and group decision-making processes amongst groups of individuals. e-knowledge networks are particularly important in the context of international development initiatives that recognize that knowledge is the key to technological change and sustainable economic development. This thesis is intended to aid knowledge network managers and researchers in their efforts towards making their knowledge networks sustainable.

The thesis addresses in depth the most important barrier towards a knowledge network's sustainability- the problem of collective action among the participants of a knowledge network. It takes the view that knowledge is a public good, and a knowledge network would face the problem of under provision of this public good due the problem of free-riding and lack of mechanisms to mobilize collective action. This thesis provides guidelines and recommendations for the restructuring of incentives and organizational policies and the re-engineering of the technology to overcome this barrier.

The thesis first outlines a framework and taxonomy for describing different knowledge network configurations and maps out the state of existence of important knowledge networks existing in the sustainable development domain within this framework. It then provides individual and comparative studies of two important knowledge networks related to sustainable development- the MIT developed Global System for Sustainable Development (GSSD) and the Development Gateway's Knowledge Network.

The engineering systems approach used in this thesis enables the study of each of the sub-systems that make up a knowledge network- the human subsystem, the technological subsystem and the institutional subsystem. This is done with the view towards providing insights into the structure of the network and the network of relationships that develop within a Knowledge Network, determining the motivations that drives the creators and the participants and the incentives that have been engineered into the technological and organizational policies to meet these motivations and assessing the quality, quantity and the evolution of knowledge and the throughput of participants in the network. A detailed description of each of the subsystems is provided and the interrelationships amongst them are analyzed and the result is synthesized to develop an integrated framework for the assessment of knowledge networks.

Thesis Supervisor: Nazli Choucri
Title: Professor, Political Science & Associate Director, Technology & Development Program

Acknowledgments

This thesis is dedicated to my wife Swati whom I cannot thank enough for her patience, love, and unwavering support. I am also thankful for the support from faculty, friends and staff in the Technology and Policy Program.

Researching and writing this thesis has been an amazing adventure and a great challenge, but none of this would have been possible without the help of countless people.

First and foremost, I would like to express my deepest gratitude to my advisor Prof. Nazli Choucri for her invaluable guidance, constant encouragement, support (intellectual, moral and financial) and personal investment into my intellectual development. Her insights, suggestions and experience have enabled me to constantly improve on my thesis and bring it to completion.

My association with Prof. Choucri and the GSSD initiative over the last year has been the inspiration and motivation behind writing this thesis and I am thankful for the kindness and support of my GSSD team members - thank you Farnaz, Toufic, Hussam, Dan and Michelle. I would particularly like to extend my appreciation to Vincent Maugis for his stimulating intellectual support and for being my bouncing board for ideas on our cigarette breaks.

I am also thankful to experts outside MIT- Rajendra Shende and Jim Curlin of UNEP, and Gerhard Pohl and Sudhakar Kaveeswar of the World Bank's Development Gateway for having me over during the summer. I would also like to thank John Maughan of the World Bank for generously sharing with me data on the Development Gateway, without which this thesis would not have been possible.

TABLE OF CONTENTS

List of Figures:

List of Tables:

1. CHAPTER I: Introduction, Problem Definition and Chapter Overview

1.1. Introduction:

The Internet and the World Wide Web have changed the nature of our social interactions. As we speak, the Web is morphing into new directions and providing new ways through which people can connect with each other. This democratization and homogenization of connections is allowing people to experience and interact with others and is drawing people together into cohesive *virtual communities* of interest.

These online communities are uniquely characterized by the strength and the depth of the ties that bind them. In Howard Rheingold's[1] words- *"...people use words on screens to exchange pleasantries and argue, engage in intellectual discourse, conduct commerce, exchange knowledge, share emotional support, make plans, brainstorm, gossip, feud, fall in love, find friends and lose them, play games, flirt, create a little high art and a lot of idle talk... People in virtual communities do just about everything people do in real life, but they leave their bodies behind. You can't kiss anybody and nobody can punch you in the nose, but a lot can happen within those boundaries. To the millions who have been drawn into it, the richness and vitality of computer- linked cultures is attractive, even addictive."*

The significance of the Internet as a powerful tool for sharing knowledge has led to an explosion in the number of *knowledge networks* on the web (e-knowledge networks). These knowledge networks combine the power of virtual communities with the collaborative tools of the Internet, enabling like-minded people sharing the same vision and values to share, access and manage tacit and explicit knowledge, online.

Electronic knowledge networks effectively facilitate the discovery, publication and application of new knowledge, the dissemination of information concerning best practices and the exchange of views and opinions amongst their users thus serving as engines for development. The constantly ongoing information exchange among networked knowledge workers, business and technology managers, scientists and engineers, government and organizational bureaucrats have greatly improved productivity, accelerated the product development cycle, revolutionized marketing and spurred innovation and entrepreneurship. Many multinational companies and international

---

[1] Rheingold, H., (1993),"The virtual community: homesteading on the electronic frontier", p3, Harper Perrenial, New York

development organizations now regard knowledge networks as a vital component of their capacity building strategy.

1.2. The emergence of e-Knowledge Networks, Virtual Communities, and Communities of Practice:

A knowledge network can be defined as the cooperation of individuals who produce, share, or use a common repository of knowledge [2]. When the knowledge network relies on the Internet as the platform for producing, sharing or using the repository of knowledge -either wholly or partly, the knowledge network is defined as an e-knowledge network.

Virtual Communities (VCs) are social aggregations that emerge from the Net when enough people carry on those public discussions long enough, with sufficient human feeling, to form webs of personal relationships in cyberspace. A Community of Practice (COP) is a self-organized, deliberate collaboration of people who share common practices, interests or aims and want to advance their knowledge on a particular subject. Thus, a COP is distinguished from a VC primarily due to its focus on the topic of discussion, and usually by the characteristics of its members – who are more often than not likely to belong to the same profession or work situation.

For this thesis, e-Knowledge Networks, virtual communities and Communities of Practice refer to the *set of Internet based tools* that support communication, collaboration and group decision-making processes amongst groups of individuals with the express purpose of sharing knowledge on a particular topic. They are distinguished from other Internet based knowledge services like portal/vortals, information clearinghouses and digital libraries primarily due to their focus on community building activities. In this thesis, e-Knowledge Networks, VCs and COPs are used interchangeably for one another.

The first virtual communities emerged about a decade after the advent of the Internet. These were the first newsgroups (USENETs) whose function was to distribute 'news' on various topics throughout the network. The growth of USENET led to the development of the BBNET in the late 1980s. The BBNET consisted of bulletin-board type discussion groups, where a participant could post comments on a news item, and read the comments of others. Since the 1980s, the growth of BBNET has been exponential- from 158 groups in 1985 to the over 40,000 groups that are

---

[2] Baets, W. R. J. Organizational Learning and Knowledge Technologies in a Dynamic Environment, p45 : KluwerAcademic Press, 1998.

currently in existence. Initially the groups were focused on technical or academic subjects, but now there are thousands of groups on a range of subjects from product user groups, to pet owners to religious groups. Clearly online networks that provide rich and authentic community experiences to their millions of users are emerging as an important social force.

Knowledge networks and meta-networks (a network of networks) are currently attracting much interest among academics, consultants and commercial and international development organizations. Academic researchers are undertaking research into how best knowledge networks can be supported and how this can help support the generation of new knowledge. There has been considerable research on the 'right' knowledge network structure (technological and organizational) to foster knowledge sharing and learning so that individual members, as well as the collective community, are more effective, in a sustainable way. Consultants and companies in the field are developing tools and techniques for supporting, coaching and facilitating these networks.

1.3. Problem Definition and Solution Approach

Although Knowledge Networks continue to grow and spread, the experiences of the last decade have shown that knowledge communities are living, open systems that have to adapt to their environment. For every successful network visible today, there have been many that have faltered and failed to realize their initial objectives as evinced by disaster stories on news websites and failed links that litter the Internet. No longer is the 'build and they will come' philosophy towards creating knowledge networks valid any more.

This thesis aims to find out why. It provides an integrated assessment framework for evaluating knowledge network performance. The knowledge networks examined in this thesis are treated as a combination of relationships, information content, work processes, technology tools and institutional architectures brought together to achieve collectively defined purposes. There little attempt to link each of these entities into a comprehensive framework to see the collective effect on a knowledge network's success, and this thesis is an attempt to fill this knowledge gap.

This thesis takes the view that knowledge is a public good, and that contributions to a common knowledge pool in the network would result in an Olsonian[3] collective action/ free rider problem. The free rider problem would lead to under provision of knowledge resources as a whole, and would undermine the knowledge network's sustainability. The critical challenge for knowledge

---

[3] Mancur Olson, The Logic of Collective Action: Public Goods and the Theory of Groups, Harvard University Press, Cambridge, Mass., 1971

networks thus is to find ways to encourage participants- who have complete discretion over how they handle their knowledge assets- to use their knowledge for the benefit of the community by sharing what they know openly and freely.

This thesis addresses the collective action problem at three levels – at the individual, network and the institutional level and provides guidelines to re-engineer the organizational processes, the technology and the institutional structure for sustained participation and co-operation within the online community. The focus of this thesis is on e-knowledge networks in the domain of sustainable development. It targets two important knowledge networks in this domain- the MIT developed Global System for Sustainable Development (GSSD) and the World Bank developed Development Gateway and by examining the characteristics of their user profiles, their processes, and their technological and institutional configurations, provides guidelines and recommendations for re-engineering these networks to make them sustainable.

1.4. Thesis Organization and Chapter Overview

Chapter 2 of the thesis focuses on the characteristics and configurations of e-knowledge networks in the domain of sustainable development. This chapter outlines the importance of knowledge in international development and touches on the complexity of mapping the domain of sustainability, which by definition requires the application of some forms of intellectual coherence.

The chapter then attempts a classification framework to classify knowledge networks according to their varying objectives, facilitating conditions, work processes, network architecture, membership structures, and institutional configurations. The chapter concludes with an evaluation framework that builds upon a two-stage model where the first level attempts to identify the *effectiveness* of a network in reaching its objectives, and the second level looks at the *efficiency* of the network in reaching these objectives.

Chapter 3 is concerned with analyzing the motivations and incentives for collaboration within a knowledge network. A literature review of the theories currently available to explain participation is first presented. The theories of Self-Interest, Exchange and Dependency theories, Cognitive theories, theories of uncertainty reduction and the theories of mutual self-interest and collective action are studied in detail and applied to the context of contribution of resources to electronic knowledge networks. The problem of collective action, free riding and the theory of critical mass are discussed in detail and their implications towards content provision are analyzed.

The second part of Chapter 3 borrows upon the experiences of three successful and well researched implementations of e-knowledge networks – corporate knowledge management systems, the Open Source movement, and the Peer-to-Peer file sharing networks all of whom have displayed a high degree of resilience and adaptability to develop mature, successful and sustainable models for participation that other knowledge networks would do well to borrow from. The chapter then identifies the motivations for cooperation and checks the extent to which the incentives engineered into these networks reinforce these motivations. The chapter concludes by offering guidelines on planning and designing e-knowledge networks at three levels – the individual, the network and the institutional level for sustained knowledge creation and sharing.

Chapter 4 & 5 are case studies of the Development Gateway's topic pages and the GSSD. An overview of the goals, user profiles, organizational processes and network architecture in each of these networks is presented. Chapter 4 contains a traffic analysis of the network and the results of this analysis towards Topic Rationalization, community building and resource allocation is discussed. Chapter 5 presents a content affinity analysis of the GSSD knowledge base and the implications of the findings towards content restructuring and improving the ontology's efficiency.

Chapter 6 discusses the findings of the preceding two chapters and attempts to compare and contrast the approach of GSSD and the Development gateway with respect to information and organizational architecture, the nature of partnerships, their technological choices and work processes. It identifies synergies between the two networks and outlines possible areas for collaborative development. Chapter 6 then synthesizes the theories, frameworks, case studies and analysis approaches from the preceding chapters to develop an integrated framework for assessing knowledge networks. It concludes with a discussion of the next generation of knowledge networks and identifies areas of further research.

2. CHAPTER II: Characteristics of Knowledge Networks

2.1. Knowledge Networks in Development:

The advent of the Internet has enabled users to accumulate and disseminate knowledge in a ways that could not have been dreamt of a decade earlier. There is an ever-increasing number of websites that provide a wealth of free information and resources on a wide range of subjects. However, the data present on the web is raw and difficult to synthesize. This inaccessible, fragmented and distributed knowledge of unknown quality has led to the proliferation of content metasites- websites that aggregate information from other websites, usually within a particular subject domain. The advent of these metasites and meta networks (collections of meta sites), herald a new phase of the Internet, one that approaches the solution to the problem of interpreting the unstructured data in the WWW.

Over the last few years, there have been quite a few domain specific portals, applications and websites with very interesting technical and business models. In the subject domain we are interested - the realm of Sustainable Development, there have been a number of organizations that have tried to create an overarching structure for analyzing the wealth of knowledge currently available (and being created every day) in the WWW. They have tried to satisfy the need for in-depth research and the need to find the best resources on specific topics within this domain. However only a few organizations have realized that to be really successful, the service an organization would have to provide cannot be achieved by technology alone. This thesis looks at two successful organizations that have tended to view technology as an enabler and how they use a combination of intelligent technologies and human specialists to serve their users better.

Our project shall focus on two successful metanetworking websites in the sustainable development domain. The GSSD (Global System for Sustainable Development)- an initiative of the Massachusetts Institute of Technology), and the DG (Development Gateway)- an initiative of the Development Gateway Foundation, and incubated by the World Bank.

Electronic knowledge networks effectively facilitate the discovery, publication and application of new knowledge, the dissemination of information concerning best practices and the exchange of views and opinions amongst policy makers and people on the field, thus serving as engines for development. The constantly ongoing information exchange among networked knowledge workers, business and technology managers, scientists and engineers, government and organizational bureaucrats have greatly improved productivity, accelerated the product development cycle, revolutionized marketing and spurred innovation and entrepreneurship in the

developing world. Many international development organizations now regard knowledge networks as a vital component of their capacity building strategy.

2.2. The Complex world of Sustainable Development

Knowledge Networks in the arena of development are dedicated to the discovery of new knowledge and its application for the advancement of developing nations and regions. The key to their popularity and initial success is the realization that all participants, both developed and developing countries and institutions, can and should learn from each other and should acquire the technology and the capacity for knowledge creation, aggregation and exchange.

The significance of the Internet as a powerful tool for sharing knowledge was emphasized by Joseph Stiglitz, former chief economist of the World Bank. Stiglitz[4] advocates as his main thesis the approach to "Scan Globally, Reinvent Locally". In other words, the global knowledge acquired from the existing repositories, such as major libraries, databases and other sources of knowledge made available on the Internet, must be internalized, rediscovered and translated to local conditions if it is to be usefully applied in development.

Sustainable development, at its heart is the simple idea of ensuring *a better quality of life for everyone, now and for generations to come*[1]. Although the idea is simple, the task is substantial. In essence, it involves meeting the four simultaneous objectives of:

- social progress which recognizes the needs of everyone;

- effective protection of the environment;

- prudent use of natural resources; and

- maintenance of high and stable levels of economic growth and employment.

"The complexity of "sustainability", by definition, requires the application of some forms of intellectual coherence – a requirement that is especially compelling in the domain of "sustainable development." In the sustainability domain, however broadly defined, the making of decisions and the formation of policy seldom draw on the full range of relevant knowledge, critical resources and overall capabilities."[5]

---

[4] Joseph Stiglitz, Scan Globally, Reinvent Locally: Knowledge Infrastructure and the Localization of Knowledge, First Development Network Conference, Bonn, Germany, December 1999. Available at http://www.gdnet.org/pdf/226_GDNfinal.pdf

[5] Nazli Choucri, et al(2003), "Mapping Sustainability", Available for download at http://gssd.mit.edu

Everyone recognizes that knowledge is the key to technological change and sustainable economic development – in all contexts and at all levels in industrial as well as industrializing countries. Despite advances in information and communication technologies, major political, strategic, economic and institutional barriers continue to impede the use of knowledge for decision-making. These reinforce the difficulties of bringing existing knowledge into the policy debates. As early as 1998, the annual Development Report, the World Bank signaled the growing dangers of increasing knowledge gaps between rich and poor nations – a problem that is now recognized as one of the most serious impediments to development. Indeed, barriers to knowledge magnify existing developmental challenges.

The explosion of information and published material on sustainable development is both a danger and an opportunity. The danger is that a vast and accelerating quantity of conceptually unstructured, difficult-to-access, and largely unevaluated material has less utility for research and policy analysis. Conversely, there is a clear opportunity that the investment of so much human capital will dramatically advance new paradigms needed for embarking on trajectories toward sustainability.

Given the complexity of the domain, there is a need for new competencies that include amongst others:

- A new information architecture that includes new languages, categories, and metaphors for identifying and accounting for skills and competencies.

- A new technical architecture that is more social, transparent, open, flexible, and respectful of the individual users.

- A new application architecture oriented toward problem-solving and representation, rather than output and transactions, and

- A new institutional architecture with organizational processes that extract the most out of the synergistic combination of IT and the creativity and innovation capacity of human beings.

Over the last decade, Knowledge Networks on sustainable development have evolved from pilot projects to large-scale programs and now towards networks that are linking up with other existing or newly formed networks. They have also grown in the number of partners and participants who now take on a wide range of roles.

In most knowledge networks, the key partner acts as a facilitator while others bring in funding, technology content and other services –like translation, hosting, organizational and administrative services etc. Some examples of these knowledge networks include – the World Links for Development, a network of connected schools around the world, the Global Knowledge Network

helping to connect researchers and policymakers, to the Global Development Learning Network, helping to disseminate knowledge to decision-makers around the world (all three networks are run by the World Bank) and the MIT run GSSD.

In later sections we shall see how each of GSSD and DG have created their own technology, editorial and institutional strategies to meet the above challenges, and the measures of success they have achieved.

2.3. Configurations Of Knowledge Networks:

There are different kinds of knowledge networks in existence, and they can be characterized along different dimensions – from the centralized to the distributed network, the global to the local, the specialized vs. general scope, from transactional systems to document repositories and content management and distribution systems. They also vary by their institutional configurations, membership structures and participation processes, the extent of moderation/quality control etc.

The term knowledge network designates both the social relationship between actors - who could be persons, groups, and even collectives of organizations, communities or even societies and the artifacts that support the relationship. Knowledge networks can thus be characterized by the form and content of their relationships, which together with the technology and organizational policy and rules establishes the network structure[6]. Networks have a structural as well as a cultural dimension and the relationships, motivations and nature of the relationship varies between network actors.

Thus, Knowledge Networks can be characterized as comprising of actors, with relationships between actors categorized by their form, content and intensity; by the resources which may be used by actors within their relationships, and institutional properties, including structural, technological and cultural dimensions such as control/governance mechanisms, standard operating-procedures, norms and rules, communication patterns, etc.

Some authors also look upon the network as a whole- as a unique hybrid[7] form of organization between market and hierarchy because they contain elements of both forms. Using von Kragh's[8] framework for characterizing knowledge networks [See Figure 1 on the next page], where the *facilitating Conditions* comprise the network's internal structural and cultural dimensions in which knowledge work processes take place, thereby defining the enabling or inhibiting environment for knowledge creation and transfer. This includes the organizational structure, the management systems and the network culture. The structural dimension includes the kind of organization the facilitator is (E.g. World Bank, MIT, Microsoft etc.) and the nature of relationships with participants (institutional arrangement between MIT and other universities, the United Nations and the National Ozone Units in different countries etc.).

---

[6] Burt, R. (1979): A structural theory of interlocking corporate directorates. In: Social Networks, 2, pp. 415-435

[7] Thorelli, (1983) . Networks: Between markets and hierarchies. Sloan Management Review, 7, pp. 37-51.

[8] Andreas Seufert, Georg von Krogh, Andrea Back(1996):"Towards Knowledge Networking", pp 5-8
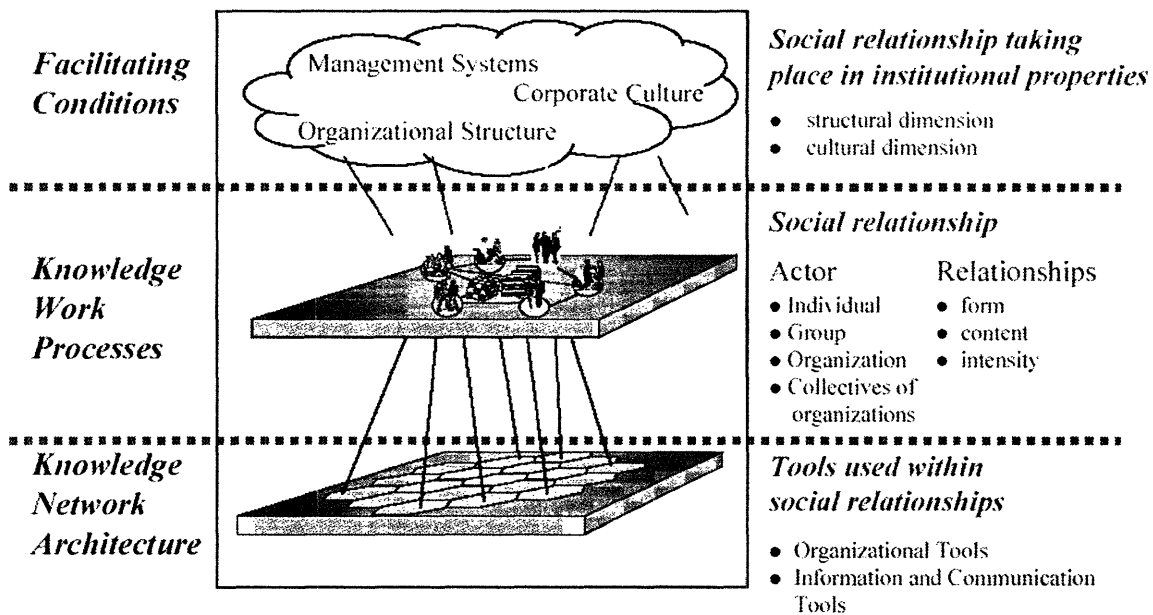
Figure 1: Analysis Framework for Knowledge Networks[9]

An example of a type of network culture is 'Care', where care as conceptualized by von Krogh (1998)[10], is crucial for knowledge creation. According to whether there is a high- or low-care environment, knowledge creation and transfer processes will differ considerably. Care involves helping behavior among people, lenience in judgment of new ideas, and an active attitude to understand others. The facilitating conditions in a network are a dynamic that varies throughout the life of a network. Both fledgling and very mature virtual communities are usually characterized by the amount of care and sympathy in the network.

The typical questions knowledge networks face are on the level of control the sponsor firm should maintain, the roles and responsibilities of the sponsor and members, the criteria for membership selection, rewards mechanisms for contributions, degree of open-ness/ hierarchy based etc.

The **Knowledge Work Processes** comprise social interaction and communication processes on an individual and at the group level. These processes can be conceptualized as an adaptation of Nonaka's[11] 4-stage knowledge spiral by viewing the processes as a transformation process between explicit and tacit knowledge from *socialization* (the exchange of tacit knowledge between individuals in order to convey personal knowledge and experience) to *joint experience* (resulting

---

[9] Andreas Seufert, Georg von Krogh, Andrea Back: "Towards Knowledge Networking", pp 15-21
[10] Von Krogh, G. (1998): Care in Knowledge Creation. In: California Management Review, 40, pp. 133.
[11] Nonaka I. (1991): The Knowledge Creating Company. In: Harvard Business Review,6, pp. 96-104.

in new shared implicit knowledge), to *externalization* (the actual exchange of knowledge between individuals and a group) to *combination* (where different fields of explicit knowledge are combined with each other to make new knowledge available on a network wide basis).

The **Knowledge Network Architecture** comprises the tool-set used within social relationships. These tools include organizational tools- includes the roles of the knowledge actors and the information and communication tools used to enable and improve knowledge work processes. They also refer to the activities of choosing the right medium and setting the context, configuration of the network, the moderation mechanisms and quality control procedures, the creation of a shared vocabulary, the nomination/election of domain experts and the technological tools towards use of lists, filters, messaging mechanisms and navigational aids.

Knowledge communities serve many different purposes and come in many different flavors, with varying degrees of boundary and criteria for membership. Some communities are focused on building explicit collections of re-usable knowledge assets while others are more focused on providing forums in which ideas and problems can be shared and discussed.

The American Productivity and Quality Center (APQC)[12] has identified and validated four primary "strategic intents" for communities that help define different community types. Most communities serve more than one purpose, but in practice, communities tend to focus on one and nurture it the most. The categories they have identified are-

Helping Communities provide a forum for community members to help each other solve everyday work problems. This type of community emphasizes connecting people across teams and organizational boundaries, and creating people-to-people connections. The World Bank developed Development Gateway, which creates a platform for practitioners in development to share knowledge, seek help and network with their counterparts around the world.

Best Practice Communities develop and disseminate best practices, guidelines and procedures for their members to apply. These communities use specific processes to validate the effectiveness and benefit of new practices. These are typically found internally within organizations, for example the Pacific Fleet Knowledge Management CoP is an example of a best practice community in the Navy. PACFLT Knowledge Managers and Leaders use the community to accelerate KM implementation and to standardize methodology throughout the Fleet to ensure the greatest impact with the least overhead.

---

[12] APQC, Successfully Implementing Knowledge Management, APQC Best-practice Report, 2000
http://openacademy.mindef.gov.sg/OpenAcademy/Central/HTML%20Folder/KM/bcp/downloads/APQC_
CoP_Report.pdf

Knowledge Stewarding Communities organize, manage and steward a body of knowledge from which community members can draw. These communities are distinguished from best practice communities in that they focus on all types of content, not just best practices, and they organize, upgrade and distribute the day-to-day knowledge their members use. These communities might also conduct research to discover new knowledge. The MIT developed GSSD is an example of this type of community.

Innovation Communities create breakthrough ideas, knowledge, and practices. The open source movement which brings together hundreds of software developers with the aim of building innovative software is the most visible example of this kind of community.

The author has had had the experience of working with two knowledge networks in the realm of sustainable development through his research assistantship at MIT at the GSSD (Global System for Sustainable Development), and his work over the summer for the World Bank's Development Gateway. We shall come back to these networks at different points in the paper to see how these networks have coped with some of the problems affecting knowledge networks in general. A brief description of these two networks and their unique characteristics are provided below-

The GSSD is an intelligent document repository consisting of thousands of 'abstracts' which link to resources (websites) on sustainable development. Abstracts can be submitted by any user (without registration), and the submitted abstracts are reviewed, translated and published by the institutional partners GSSD works with. The institutional partners are educational institutions belonging to different countries and they assume the responsibility for abstract review and translation. The challenge and the focus within GSSD has been to selectively aggregate content while maintaining a balance in the quality control process. This is very important because a lax selection process may lead to user searches yielding a disproportionate number of irrelevant results, and a strict quality control process may result in under populated repositories that yield too few results.

Document repositories are a unique type of knowledge network as the process of submission of knowledge is largely unidirectional and impersonal[13]. Unlike more collaborative knowledge network which typically involve N⇔N interactions amongst the actors of the network, the contribution into a document repository is viewed as an impersonal interaction with a machine and individual decisions to contribute information are based on the perception that the particular

---

[13] Subramani, M.R. and Peddibhotla, N, "Contributing to Document Repositories - An Examination of Prosocial Behavior" pp 4-6, Carlson School of Management, University of Minnesota, Undated

item of information may be useful to others. Another unique phenomenon in document repositories is that the extrinsic rewards for contribution are often nonexistent and there are no direct incentives for contribution. Another distinguishing character of document repositories is that contributors receive very little feedback on their submissions. In GSSD, the only feedback the contributor receives is that his abstract has been acknowledged by the system, and there is a chance that it is reviewed and published. Document repositories have also been characterized as altruistic networks[14] in the sociological literature as the motivations and incentives for people to contribute are not readily apparent.

The other knowledge network the author was involved with - the Development Gateway is focused more on collaboration and knowledge sharing features and is more akin to the conventional knowledge networks we discussed before. The DG offers virtual communities centered on various development issues (from Afghanistan reconstruction to 'Youth for Development') and encourages collaboration amongst development practitioners. Expert guides and global topic advisors work with their communities to highlight the most relevant and useful resources available on the Internet. Registered members can contribute resources, comment on existing resources and discuss issues of interest within their topics. The focus of the Development Gateway has largely been towards creating a sense of community to the members participating under the various topics.

---

[14] Desouza, K.C., "Socially Altruistic Electronic Knowledge Communities: Avenue for Research and Preliminary Insights," In Proceedings of the Fifth Annual Conference of the Southern Association for Information Systems, Savannah, Georgia, (Mar. 1-2, 2002).

2.4. Measurement and Evaluation of Knowledge Networks

As the preceding section has shown, there are different types of knowledge networks – task oriented networks, capacity building networks, operational networks, research networks, and so on, with unique technological and organizational configurations, and with each serving the different needs of their members. Whatever the design configuration, the ultimate purpose of the network should be its capacity to provide what can be called as the 'network advantage'- that the network is more than the sum of its parts and not just an arrangement to link discrete units, or facilitate associated databases. The 'network advantage' helps make a case that operating in a network mode does lead to focused collaboration, better informed research results, new knowledge and real influence[15]. This thesis examines if this network advantage has truly been achieved by the Development Gateway and the GSSD and if not, seeks to identify the barriers that prevent these two networks from realizing this advantage.

The last decade has seen a high failure rate amongst web based knowledge networks. Failures come at a significant cost, but it is important to focus on them as they can tell us a lot, if not more than success stories. The study of knowledge networks is a new discipline of research and it is not clear why certain networks succeed and why others fail. Ultimately it is the "users" of a network who should decide whether a network should continue. Institutional Knowledge Networks have the potential to weaken the capacity of participating institutions, if energies devoted to institutional priorities and to network priorities are unbalanced.

One oft cited reason for high failure networks in the last decade has been that organizations have focused more on tools and technology of the network than on content, organizational culture and motivational approaches needed for creating a sustainable network. Effective knowledge networking needs a supportive, collaborative culture that has to be sustained at all stages in the evolution of the network. Another frequently cited reason is that the knowledge provided is poor in terms of consistency and relevancy. This is a direct result from the third most often-cited reason -the absence of a feeling of 'community' or the presence of a hostile/fragmented one.

Other factors – like the lack of adequate financial support and network 'champions', the choice of incorrect technology, organizational processes and/or governance mechanisms also play a significant role and each of these factors are discussed in subsequent sections. This section is concerned with the periodic measurement and evaluation that could help prevent or at least diagnose these factors before they lead to failures.

---

[15] Heather Creech and Terri Willard (2003): "Strategic Intentions: Managing knowledge networks for sustainable development" pp 23-30, IISD Working Paper: International Institute for Sustainable Development.

The focus of this thesis is on re-engineering knowledge networks to help them achieve sustainability, sustainability here defined by the US Navy's [16]use of the term *"The attributes and status of a system and its supporting logistics infrastructure that enable an operational system to "maintain" its mission capability during the course of operations".* However, the previous definition of sustainability is conspicuous by its absence to include the adaptability of the network to its environment. Our definition conflates the above definition to include this characteristics - in other words sustainability here is the ability of the network to carry out its stated mission and goals consistently over time by adapting itself to the needs of its environment.

In their paper[17] Conklin et al build a theory on the ecology of sustainable knowledge networks where they say that " ..by focusing on sustainability, we do not mean to imply that knowledge networks should last forever, or that any pause in growth implies an ecological problem. .. Knowledge networks have a natural rhythm and pace to them; before they die they may go to sleep and even go dormant for periods of time. The practical challenge for knowledge networks is with illness and potentially early death – the failure of the ecology to sustain a virtual team for its chartered or intended life span."

Clark in his survey of Canada's evolving knowledge networks[18] states that the main purpose of a network is to create and disseminate knowledge for use beyond the membership of the network, with the structure and operation of the network designed to maximize the rate of knowledge creation. He lists the following "ideal characteristics" of formal knowledge networks:

> ➢ Provides recognizable direct benefits to participants.
> ➢ A formal organization with a well-defined management structure.
> ➢ Participation is by invitation, based on criteria of merit or peer review.
> ➢ There is a well-developed communications strategy; and
> ➢ The network results in a reduction of boundaries between sectors, such as universities and industry.

He also identifies additional elements of a successful knowledge network including among others:
> ➢ culture shifts within institutions towards collaborative activities between institutions and sectors;

---

[16] US Navy's glossary of terms for Contractors: http://www.acq-ref.navy.mil/tools/turbo/topics/cp.html
[17] Conklin, C. Ellis, L. Offermann, S. Poltrock, A. Selvin, J. Grudin (1998), Towards an Ecological Theory of Sustainable Knowledge Networks, GDSS Inc. Available at www.touchstone.com/tr/wp/ecology.html
[18] Clark, H.C. 1998. "Formal Knowledge Networks: A Study of Canadian Experiences." Winnipeg, Canada: International Institute for Sustainable Development.

> multidisciplinary, multisectoral and multi-national/regional in terms of both network participants and in audience;

> cost effectiveness in operations, and possibly revenue generating through sales of products; and frequently mobilization and/or more efficient use of human resources;

> more effective influence on decision-makers through size of network, reputation of network members and quality of collaborative work – this is maintained through a careful balance of management and degree of selectivity

Traditional frameworks and models used to evaluate the effectiveness of a knowledge network have focused on the operational elements of running a network. The IISD (the International Institute of Sustainable Development) - currently engaged in research on Managing Knowledge Networks for sustainable development, advocates a more holistic view. Specifically, it takes the position that a network needs to be able to determine what changes it has effected through its research and communications work. It needs to monitor whether it is fully realizing its potential. This requires evaluation methods that not only assess individual activities, but also provide some means for identifying changes as a result of its combination of efforts[19].

Their evaluation takes a hierarchical approach to, where effectiveness –towards reaching the overall goals of the network is first evaluated, and then the efficiency of the network in reaching those goals is monitored on a periodic basis. This is similar to the Logical Framework Analysis (LFA) [20]approach used by development practitioners to plan and monitor projects. The core elements of the LFA involve stating explicitly the goals, purposes, outputs and activities of the network and devising indicators, means of verification and underlying assumptions and risks of the measurement process. This hierarchy also maps directly onto Figure I, where the effectiveness of the activities in meeting their stated objectives are compared along the dimensions of network architecture, knowledge work processes and the extent to which conditions facilitate meeting the network's objectives

The effectiveness of the network (doing the right thing) is defined as the ultimate value of a knowledge network to its users is in joint value creation, building capacity and providing the information tools to influence policy processes. In this context, measuring the effectiveness of a network would ultimately involve measuring to some extent the value it has ultimately provided its end users. More specifically, if the network has explicitly stated that its purpose is to lever change in policies and practices, supportive of sustainable development it needs to be able to determine

[19] Heather Creech and Terri Willard (2003): Strategic Intentions: Managing knowledge networks for sustainable development, IISD, Chapter 6, pp 117
[20] United Nations Development Programme (UNDP): Global Environment Facility. Information Kit on Monitoring and Evaluation. 1999.

what changes it has effected through its research and communications work on its users. It needs to monitor whether it is fully realizing its "network advantage". This requires a methodology that not only assesses individual activities, but also provides some means for identifying changes as a result of its combination of efforts.

Measuring the *effectiveness of the network* involves defining the overarching goals and the ultimate purpose of the knowledge network. Although most knowledge networks state their goals upfront, it is important that these goals are well defined, clear and are endorsed by its members.

An example a well defined goal is that of the Climate Change Knowledge Network's mission: '.. to undertake collaborative research and action on issues such as the Kyoto Negotiators; mechanisms, adaptation Vulnerability and to climate change and Adaptation; technology transfer; and to build capacity in developing and developed countries to better understand and address climate change issues; and to communicate information and knowledge within and outside the network."[21]

The overarching goal should lead to answers on questions like what can members contribute to, as well as receive from, the network; what will success look like for the network as a whole; who is going to benefit, be changed or influenced by the work; and what will be the indicators of success for each activity?

Once a clear overarching goal is defined and the above questions are identified, the next stage is to identify objective metrics that measure performance towards that goal. If the goal were to communicate information and knowledge within and outside the network a simple metric like the number of articles/documents downloaded/uploaded would suffice. If the goal is to build capacity by collaboration and through the network advantage, the measurement becomes much more trickier. Evidence of member's strengthening each other's work/knowledge can be tracked by their participation in discussion groups, number of research papers circulated for comments/peer review, evidences of co authoring etc. If the goal of the network is to influence policy and processes, the measurement could be very difficult without in-depth surveys and interviews. Proxy indicators like members participating in events or hosting workshops or securing face-to-face meetings with decision makers they would otherwise have not had access to.

Once the overarching goal is agreed upon and communicated, the work plan and the outputs to meet the objective metrics defined in the previous stage could be created, the next level of measurement involves evaluating the *efficiency of the network* (doing things right). This

---

[21] From the Climate Change Knowledge Network (CCKN) website : http://www.cckn.net/about.asp

approach focuses on the transactional costs of setting up and running networks that they are cumbersome and time-consuming to manage, where motivation and performance of individual members is often at issue and that the cost effectiveness of the network approach is in question. The aim of this exercise is often to monitor, review and adjust the internal operations of the network on a periodic basis.

Traditional literature on evaluating the operational performance of knowledge networks tends to take a ROI (return on investment) approach to the costs and benefits of investing in initiatives with in the network. This approach when applied to knowledge networks on sustainable development may defeat the very purpose the knowledge network was built for- provision of the public good to people in developing countries. Non-profit organizations have unique needs and requirements that differentiate them from businesses. This however does not mean that it is incorrect to focus on operational parameters or on corporate knowledge management evaluation systems like cost benchmarking and balanced scorecard methods to improve efficiency.

This thesis uses two different modes of operational analysis for gauging a network's path towards sustainability.

The first approach of analysis is concerned with the relevance, usefulness and the accessibility of the content within the network and takes the view that the value of content is based upon the combination of its primary useable form, along with its application, accessibility, usage, usefulness, and uniqueness. Content analysis aims to understand and match content to user needs with the aim of facilitating collaboration and knowledge sharing through greater accessibility of the content.

The content analysis approach inventories the knowledge resources available within the network, identifies the characteristics of content provision, and seeks to improve upon the organization of this content to build value in the network by refining the set of rules and processes for contributing, collaborating on and controlling content. One of the aims of the content approach is in refining the ontology for displaying content. The content analysis exercise of the GSSD Knowledge Base described in section 5.2.1 explains this approach in greater detail

The second approach to knowledge network evaluation is more concerned with the usage and traffic to the network. One of the advantages of electronic knowledge networks is the ease with which data can be gathered about their usage. All websites generate log files that can be analyzed by a wide variety of software tools or outsourced services. The difficulty lies in deciding

precisely what to analyze, and in determining the real significance of the data. Are 'page impressions' a more accurate metric than 'hits', for example, or should we forget them both and concentrate on 'conversion rates' and 'repeat visits'. A detailed traffic analysis of the Development Gateway portal outlined in Section 11.2 explains this approach in greater detail. Briefly traffic analysis is concerned with -

- The *Size* of the online community (both unregistered and registered users) and the size of the message base (absolute and relative to the expected population)
- *Growth rate*- Number of new members voluntarily adding subscriptions vs. defections.
- Increasing *number of contributors* (as a percentage of subscribers, although this rate is always low. One study found that half the messages in successful groups came from two percent of the members.)
- *Relative activity:* total postings, postings/contributor, growth in postings, increases in thread length.
- The *relative number of "sanctioning" messages,* from comments to "flames" required keeping contributions on topic. No sanctioning messages might indicate a dead community (no one cares any longer), while too many sanctioning messages might indicate a discourse community having difficulty establishing the proper norms and expectations.

Both content and traffic analysis provide useful tools to gauge the health and robustness of the network and need planned data collection mechanisms and organizational processes to make sure they are carried out on a periodic basis. These analysis tools provide an objective view of the performance of the knowledge network. To assess aspects of organizational culture and the extent to which people's opinions, attitudes and behaviors are changing these studies could be complemented by user surveys. Although surveys measure people's subjective perceptions and these may or may not reflect reality, that can in many ways be their very benefit, as people's perceptions will determine their behaviors with respect to knowledge collaboration and sharing.

3. CHAPTER III: Motivations and Incentives for participation in Knowledge Networks

This thesis aims to provide solutions that increase knowledge-workers' effectiveness by helping them share knowledge and learn through participation in knowledge networks. Sharing has been identified as a critical cultural change and the largest single stumbling block to the success of knowledge networks. Many networks have invested in expensive architecture, created roles, appointed people and offered incentives but have not been able to build a sharing ethos.

This chapter explores the motivations, culture, rewards, structures, styles and tools for sharing knowledge. The first part of this chapter looks at the existing literature on knowledge networks and by examining trust, communication styles, behaviors, social rituals, group dynamics, the role of identity and the identification of sharing patterns, attempts to map the terrain of theories related to explaining the motivations and incentives for sharing knowledge.

The second part of this chapter looks at three kinds of implementations of knowledge networks and attempts to understand the specific motivations of individuals participating in these network and use this learning for re-engineering knowledge networks on development. These three networks – corporate knowledge management systems, the Open Source community and the Peer to Peer file sharing systems have displayed a high degree of resilience and adaptability to develop mature, successful and sustainable models for participation that other knowledge networks would do well to borrow from. This chapter concludes with three categories of solutions for re-engineering- solutions geared to restructuring individual incentives, solutions that attempt to change the characteristics of the network, and solutions that involve reconfiguring the institutional structure.

3.1. Theories of Knowledge Sharing-

Before studying the motivations for participation within knowledge networks, it would be useful to look at the evolution of the notion of knowledge itself in the economic, sociological and organizational science domain.

The literature on Knowledge is characterized by three distinct perspectives – those that view knowledge as an object, those that view knowledge as something embedded within individuals, and the recent publications that view knowledge as embedded in a community[22].

---

[22] M. McLure Wasko, S. Faraj1(2000): "It is what one does": why people participate and help others in electronic communities of practice, Journal of Strategic Information Systems 9 (2000) 155-173

The *knowledge as an object* view assumes that knowledge can exist independent of human action and perception. It is conceived as some knowable truth that can be codified and separated from the people that posses it. By adopting this perspective, the goal of knowledge management is then to convert the knowledge residing in the minds of people, into structural assets owned by the firm and store it in the firm's knowledge management system (KMS).

The *knowledge embedded in individuals* perspective views knowledge as inseparable from people, suggesting that knowledge resides only in the minds of individuals. Only people can 'know' and convert 'knowing' into action, and it is the act of thinking that can transform information into knowledge and create new knowledge. In addition, people seem to know a great deal more than they can articulate and this tacit component of knowledge has a personal quality which makes it hard to formalize and communicate. Thus this perspective focuses on the management of human resources. Since this perspective views knowledge as difficult to codify and suggests that knowledge loses its value once codified, the goal of KMS is to connect experts with knowledge seekers.

The *knowledge embedded in a community* perspective defines knowledge as 'the social practice of knowing' and emphasizes that learning, knowing and innovating are closely related forms of human activity and inexorably connected to practice. Each community develops its own language, shared narratives and codes, and knowledge is best understood within the context of its community. This view attempts to locate organizational knowledge and knowledge creation within distributed, multi-actor routines rather than individual minds and the focus is on processes that are geared towards enabling discussion, mutual engagement, and exchange between members of a community.

The knowledge as object and knowledge embedded in people perspectives view knowledge as a *private good* (Private goods are goods with high excludability and high rivalry. It is clear who benefits from the good or service and that person/organization can easily be charged), owned either by the organization or the individual. In such cases, people exchange their knowledge through market mechanisms in order to receive commensurate benefits. They are motivated by self-interest and are less likely to exchange knowledge unless provided with tangible returns such as promotions, raises, and/or bonuses, or intangible returns such as reputation, status and direct obligation from the knowledge seeker.

The knowledge embedded in a community view takes the stand that knowledge is a *public good* (A public good is one where users collectively consume benefits and no one can be excluded

from consuming the good), and that the economically rational action is to free-ride, in other words, consume the public good without contributing to its creation or development. The motivation to maximize self-interest does not adequately explain why people contribute to public goods when it is not rational to do so. We shall discuss the implications of treating knowledge as a public good and the theories of self-interest and collective action in the next section.

In the sociological view, there has been an evolution from a focus on the traditional epistemology of knowledge –concerned with the understanding the origin, nature and validity of knowledge, towards a social view which looks at knowledge as the output of the interplay between individual contributions and social interactions. The literature related to the characterization of knowledge in sociology has largely been along the following lines[23] –

- As a social construction, embedded in a system of social relationships
- As being socially spread, and influenced by social settings
- As something that is developed through participations in communities of practice
- As an epistemological construct that is continuously changing: from individual to social, from tacit to explicit

The implications of the sociological view for knowledge creation are that knowledge creation is seen as an emergent process (and not the output of a rigid planning process), or as a shared outcome of individual actions, where every player contributes, but individual contributions cannot be uniquely identified. The outcome is that knowledge is not isolated and owned by individuals, but is distributed and owned by a community of individuals.

On Knowledge sharing and knowledge networks, there is large amount of literature from researchers in the sociology, organizational behavior, economics and mathematics (chaos theorists) domains. Monge and Contractor[24] have presented an excellent overview of the underlying theories and the mechanisms to explain the emergence of knowledge networks in their book. Some of the more salient theories are presented below.

---

[23] Sawhney, Mohanbir and Prandelli, Emanuela, "Communities of Creation: Managing Distributed Innovation in Turbulent Markets", California Management Review, Vol. 42, No. 4, 2000, pp. 24-54
[24] Monge, P. R., & Contractor, N. S. (2001): Emergence of communication networks. In L. Putnam & F. Jablin (Eds.) New handbook of organizational communication. Newbury Park, CA: Sage.

Theories	Theoretical Mechanisms
Theories of self-interest Theory of Social capital Transaction Cost Analysis	 Investments in opportunities Minimize the cost of transactions
Exchange and Dependency theories Social Exchange Theory Resource Dependency Theory	 Exchange of valued resources (material or information)
Cognitive theories Semantic Networks Knowledge Structures Cognitive Social Structures Cognitive Consistency	Cognitive mechanisms leading to: Shared interpretations Knowledge transfer Similarity in perceptual structures Drive to restore balance
Theories of uncertainty reduction Uncertainty reduction theory, Contingency theory	 Reduce uncertainty by communicating
Theories of mutual self-interest and collective action Public Goods Theory Critical Mass Theory	 Inducements to contribute Number of people with resources and interests

Table 1: Theories to explain Knowledge Network Participation[25]

3.1.1.   Theories of Self Interest

The evolution of knowledge networks seems a natural extension of the theory of self-interest. To illustrate, the theory of Social Capital, popularized by Burt's (1992)[26] 'theory of structural holes', argues that people accumulate social 'capital' which they invest in social opportunities from which they expect to profit. It assumes that people are motivated by self interest, and the accumulation of social capital is accompanied by an identification of network 'holes' which people seek to exploit or 'fill'. More and more people are relying on their knowledge of communication networks to accumulate capital and fill holes, but it is still unclear whether new technologies like the e-mail and the Internet make it more easier or difficult to asses the existing social structure.

The transaction cost analysis (TCA) theories, provides a macro perspective on the theory of self-interest. It is used by economists and some organizational theorists to propose that a shift from markets to hierarchies will reduce transaction costs in situations characterized by uncertainty, bounded rationality and opportunism[27]. *This proposition is exactly the reverse of the trend we are seeing, where traditional hierarchical systems are giving way to network based knowledge exchanges.*, leading one to conclude that the TCA perspective is limited as it focuses on the benefits to an individual firm and environments where the process of knowledge sharing is needed to overcome individual cognitive limits. The focus is more towards minimizing transaction costs incurred by individual firms, and this approach is now being modified[28] to maximize transactional value created by a network of firms, reflecting the fact that network strategies may have more value, or are more effective than traditional hierarchy based knowledge sharing mechanisms.

3.1.2.   Exchange and dependency theories

Exchange and dependency mechanisms have also been used to explain the emergence of networks. One of the main theories under this category is the Social Exchange Theory - developed by Homans[29], which seeks to explain human behavior through the exchange of material and information resources. From its original avatar of explaining the dyadic exchanges within organizations, it has evolved into Network Exchange[30] theories that postulate that an

---

[26] Burt, R. S. (1992). Structural holes: The social structure of competition. Cambridge, MA: Harvard University Press.
[27] O.E Williamson(1985): The Economic Institutions of Capitalism, New York, NY: The free Press 1985
[28] E.zajac and C.Olsen(1993): From Transaction costs to Transactional Value Analysis: Implications for the study of interorganizational strategies. Journal of Management Studies, 30 (1993) pp 131-145
[29] Homans, G.C. (1950). The human group. New York: Harcourt Brace.
[30] Bienenstock, E.J., & Bonacich, P. (1992). The core as solution to exclusionary networks. Social Networks, 14, pp 231-244.

individual's power to bargain is a function of the extent to which he is vulnerable to exclusion from communication and other exchanges within the network.

The argument is that individuals forge network links based on their analysis of the relative costs and returns on investments. Likewise, individuals maintain links based on the frequency, the uncertainty, and the continuing investments to sustain the interaction. This theory has been applied to study the exchange mechanisms in power, leadership and trust and ethical behavior. As people and organizations find their exchanges no longer rewarding or as new or competitive others offer better bargains in the exchange, linkages begin to dissolve.

Nahapiet & Ghoshal (1998) [31] extend social exchange theory to create the concept of a knowledge markets, where knowledge sellers work out whether it is worth sharing their knowledge with a knowledge buyer and buyers work out whether they are able to offer something in exchange such as help in the future. In this knowledge market exchanges make social capital as well as intellectual capital: "social capital is created and sustained through exchange... social capital facilitates exchange"

The resource dependency theory looks at the organization's response to changes in the organization's environment. Organizations' response to changes is to either increase the number of exchange alternatives by creating new network links, or to decrease the number of exchange alternatives for others by forming a coalition with other resource providers. These counterbalancing mechanisms provide an explanation for the stability of exchange relationships and potential redistribution of power among the individuals. An important finding with respect to the network performance was carried out by Provan[32] who argues in his paper that the network structure and the network's effectiveness is influenced by the existence of a relatively munificent environment and the degree to which the overall network is stable.

3.1.3. Cognitive theories

Yet another viewpoint is that of cognitive psychologists, who contend that any process of knowledge socialization and collective learning is based on relationships of meaning building and sharing, and as such relationships cannot be enacted in the absence of a context of co-participation, it is important to realize a minimum common denominator for all the individuals and groups participating in knowledge creation. They are concerned with the processes that promote the development of shared values, reciprocity and mutual trust.

---

[31] Nahapiet, J., & Ghoshal, S. (1998). Social capital, intellectual capital, and the organizational advantage. Academy of Management Review, 23(2), pp 242-266.
[32] Provan, K.G. (1983). The federation as an interorganizational linkage network, Academy of Management Review, 8, pp 79-89.

In this context, the OB practitioners have studied the underlying motivations for knowledge sharing, and the incentives for encouraging participation. This line of research is characterized by its recognition of the importance of the 'edges' – or network nodes, relative to the 'center' of the network. This line of research is concerned with creating an effective environment for new knowledge creation in an environment where knowledge is increasingly distributed and tacit.

The cognitive theories explain attitudes and behavior based on individuals' actual interactions. They categorized individual cognitions through–

Semantic Networks: The focus is on message content, and the shared meanings of individuals in a network on the message content. Linkages are created between people who have similar interpretations of the message.

Knowledge Structures: The focus here is on understanding why interorganizational networks are created. Interorganzational networks are viewed as structures of knowledge, and relationships between organizations are formed after a search process that takes into account the skills, competencies, trustworthiness and capabilities of other organizations.

This theory also posits that once organizations choose partners, however, they tend to spend less time seeking other partners. As Kogut, Shan, and Walker (1993)[33] say, *"because information is determined by previous relations and in turn influences the subsequent propensity to do more relations, the structure of the network tends to replicate itself over time. The early history of cooperation tends to lock in subsequent cooperation"*. This is an important lesson for practitioners who are setting up their knowledge networks and evaluating collaboration partners. They also observe that *"The replication of the network is a statement of the tendency of learning to decline with time. The structure of the network is a limiting constraint on how much new learning can be achieved."*

Both the Cognitive Social Structures and the Cognitive Consistency theories[34],[35] focus on member's cognitions, and the member's perceptions of social structures, and attempts to

---

[33] Kogut, B., Shan, W., & Walker, G. (1993). Knowledge in the network and the network as knowledge: Structuring of new industries. In G. Grabher (Ed.), *The embedded firm: On the socioeconomics of industrial networks.* New York, NY: Routledge. pp. 67-94

[34] Corman, S. R. (1990). A mode of perceived communication in collective networks. Human Communication Research, 16, pp 582-602.

[35] Freeman, L.C. (1992). Filling in the blanks: A theory of cognitive categories and the structure of social affiliation. Social Psychology Quarterly, 55, pp 118-127.

reconcile this with actual data on communications and network data. The research on cognitive social structures when applied to electronic knowledge networks is faced with the same problem the structural hole theorists face, which is determining if the pervasiveness of electronic communication media in virtual organizations makes more or less difficult for individuals to discern social structures.

3.1.4. Theories of Uncertainty Reduction

Uncertainty reduction and contingency theories deal with explaining the impact of uncertainty in individual and organizational environments. The intersection of contingency/uncertainty reduction theories with network literature seeks to explain how people and organizations use communication networks to reduce their uncertainty. Albrecht and Hall (1991)[36] found evidence that the need to reduce uncertainty also explained the creation of dominant elites and coalitions in innovation networks, and at the interorganizational level, Granovetter (1985)[37] argued that organizational decision makers use social networks to reduce uncertainty associated with market exchanges, thereby reducing their transaction costs – similar to the conclusion of the Transaction Cost Economists. Shrader, Lincoln, and Hoffman (1989)[38] using contingency theory, argue that organic forms of organizational structure would result in informal organizational communication networks that were denser, more highly connected, and more multiplex than those found in mechanistic/hierarchical organizations. Both theories have examined the impact of network structures on uncertainty reduction, but work still needs to be done to investigate the impact of the network on equivocality reduction(Weick 1979)[39].

3.1.5. Theories of Mutual Self Interest and Collective Action

Economists have been concerned with the question of allocating economic rents from the knowledge sharing process in the network and the related question of the incentives to ensure continuous creation and distribution of knowledge. This view takes the view that knowledge is a private good. However, a unique characteristic of e-knowledge networks where knowledge is created collaboratively is that the value of the network lies in giving its knowledge away freely.

---

[36] Albrecht, T. L., & Hall, B. (1991). Relational and content differences between elites and outsiders in innovation networks. *Human Communication Research, 17*, pp 535-561.
[37] Granovetter, M. S. (1985). Economic action and social structure: The problem of embeddedness. *American Journal of Sociology, 91*, pp 481-510
[38] Shrader, C. B., Lincoln, J. R., & Hoffman, A. N. (1989). The network structures of organizations: Effects of task contingencies and distributional form. Human Relations, 42, pp 43-66.
[39] Weick, K.E. (1979). The social psychology of organizing. (2nd ed.). Reading, MA: Addison-Wesley.

This line of thought views knowledge as a public good. The Linux OS was created by a community of creation that owns and maintains the Intellectual Property and gives it away for free. Economists are looking towards new mechanisms to allocate the rents in community centric innovation models and are struggling with tradeoffs between more production of ideas (by offering incentives to individuals like patents, exclusivity contracts etc.) and the faster and widespread distribution of these ideas.

Finally, some theorists have looked at the application of Public Goods Theory(PGT) to explain how people could be induced to contribute to the knowledge pool in knowledge networks. Its main focus is on the benefits of coordinated action(and mutual self interest) rather than on individual self-interest. The remaining part of this section will outline the salient aspects of PGT in greater detail and look at the implications of some of its results towards re-engineering the norms and incentives in knowledge networks.

3.1.6.   Introducing Public Goods Theory:

The Public Goods Theory(PGT) was brought about to explain and solve the public good dilemma. A public good constitutes a shared resource from which every member of a group may benefit regardless of whether or not they personally contribute to its provision, and whose availability does not diminish with use (Olson 1965)[40].

Public Goods have two distinct characteristics, they are *non-rivalrous*- several individuals can consume the same good without diminishing its value and they are *non-excludable* - an individual cannot be prevented from consuming the good. A public park is an example of a public good. We enjoy the surroundings of a park (independent of whether we pay taxes or not) and our enjoyment of it does not diminish its availability to others.

The public good problem arises from the fact that since access to a public good is not restricted to contributors only, there is a temptation for individuals to free-ride, i.e. to enjoy the resource without contributing to its provision. The dilemma is that most people in a public good situation would be happier enjoying the good at the cost of their individual contribution than not enjoying the good and saving that cost. If there were an assurance that everybody else was going to pay his or her share, most people would happily contribute as well.

To reformulate the dilemma in the language of Game Theory- Not cooperating yields the best individual utility (dominant strategy) regardless of what everyone else in the group does because-

---

[40]   Olson, Mancur (1965) "The logic of collective action" Cambridge, MA: Harvard University Press.

If everyone else cooperates and I do not, I enjoy the good for free. If no one else or very few others cooperate, I will be saving a wasted contribution. Thus if everyone acted 'rationally' according to the dominant strategy, no one would cooperate, and everyone would end up suffering the consequences. This situation is usually referred to as a deficient equilibrium. It is an equilibrium situation because, once it is reached, there is no individual incentive to break it. It is deficient because the entire group suffers a non-optimal outcome. In a sense, once the deficient equilibrium is reached, there appears to be a *social fence* that keeps the group from reaching the benefits of cooperation (Messick and Brewer)[41].

Individual	*Collective*	
	Cooperate	Defect
Cooperate	1,1	-2, 2
Defect	2, -2	-1, -1

**Payoffs in an individual vs.
collective scenario**

Table 2: Payoff matrix in Prisoner's dilemma

Hardin [42] compares this dilemma with that of the Prisoner's dilemma situation in Game Theory, where the game is played between an Individual and a Collective. The payoffs are structured to reflect the fact that if an individual and the collective decide to co-operate then all participants benefit. When either the individual or the collective contributes, one of them is at a loss whereas the other benefits without having contributed. When the individual and the collective both defect then both the actors gain nothing. The challenge, as will be seen in subsequent sections, is to use selective incentives restructure the payoff matrix to encourage cooperation.

3.1.7.   The concept of Knowledge as a Public Good:

Knowledge satisfies the two requirements we outlined above for a public good. Knowledge is *non-rivalrous-* there is a zero marginal cost from an additional individual enjoying the benefits of the knowledge and it is *non-excludable.* The fact that to acquire and use knowledge individuals may have to expend resources and that there may be significant costs associated with transmission of knowledge does not in any way affect the public good nature of knowledge itself: The good itself remains free.

While the non-rivalrous property of a public good says that no one *should* be excluded from the enjoyment of a public good, the non-excludability condition implies that no one *can* be excluded. This means that knowledge cannot be provided privately.

---

[41]   Messick, David M. and Marilynn B. Brewer(1983): "Solving social dilemmas: A review" Review of personality and social psychology. L. Wheeler and P. Shaver (eds.), pp 11-44. Beverly Hills: Sage.
[42]   Hardin, R. (1982). Collective Action. Washington, Resources for the Future.

Some forms of knowledge are excludable - trade secrets, copyrights and patents can all be used in varying degrees to exclude others from gaining knowledge, and hence in the economics literature knowledge is often characterized as an impure public good. However, in the context of knowledge networks where participation is defined as the contribution of knowledge not governed by intellectual property mechanisms, knowledge can be viewed as a pure public good.

Joseph Stiglitz[43], the ex-chief economist of the World Bank makes a convincing case for knowledge as a *global* public good. He says that the general theory of Public Goods articulated by Samuelson[44] recognized that the benefits of some public goods were limited geographically. These were called local public goods. Stiglitz argues for the existence of certain global public goods that accrue to everyone in the world. He identifies five such global public goods: international economic stability, international security, the international environment, international humanitarian assistance, and knowledge.

On Knowledge, he argues that "..although there are some kinds of knowledge which are of value only or mostly to those living in a country (e.g. knowledge particular to a country's institutions, weather, geography), scientific truths- including many of the propositions of the social sciences- are universal in nature..". He goes on to say that "..much of the knowledge that is required for successful development is not patentable as it is not the knowledge that underlies new products or new processes but the fundamental knowledge on how to organize firms, how to organize societies, how to live healthier lives in ways which support the environment, how to design of economic policies that promote economic growth etc. This knowledge which includes best practices, the accumulation of successful anecdotes and the analysis of why certain policies and practices work in some circumstances and not others etc.,   needs to be efficiently produced and equitably distributed."

Given the importance of knowledge to society, the issue of efficient production and equitable use of global knowledge is a very pressing and urgent one. However, the challenge facing knowledge networks is in the creation of efficient systems for knowledge production and dissemination in the face of free riding. The next four sections deal with solutions to this problem.

---

[43] Joseph Stiglitz(1999) "Scan Globally, Reinvent Locally: Knowledge Infrastructure and the Localization of Knowledge" Keynote Address-First Global Development Network Conference December 1999; Bonn, Germany

[44] Samuelson, Paul. 1954."The Pure Theory of Public Expenditure," Review of Economics and Statistics, 36, 4, pp.387-9.

3.2. Solutions to the collective action problem:

Given that the efficient production and distribution of knowledge requires collective action, how do we solve the problems of free riding and knowledge hoarding of actors who take prefer to take more out of the knowledge commons than they contribute. Answering these questions involves understanding the puzzle of how the groups we observe in our everyday lives overcome this free-rider problem as there is plenty of evidence to prove that the provision of collective goods is, in many cases, above the Nash prediction of the prisoner's dilemma model illustrated above.

Answers to this puzzle have ranged from using the concept of 'Peer pressure' to explain the need for individuals to comply and contribute their fair share to the evolution of institutional actors who create suitable rewards and sanctions for participations. The original formulations treated individuals as if they were isolated and independent of others making similar decisions. Oliver and Marwell [45](1993) have criticized this view and emphasized the importance of the network of relations in which people are embedded. They show that the extent to which people are interconnected in communication networks increases their willingness to support the collective good. However, each of these answers provide only partial solutions to the collective action problem. This thesis develops an integrated framework to provide a holistic solution that addresses the problem at three different levels-

At the Individual Level

• *Through the use of selective incentives:* In this approach, the free rider problem is overcome providing a selective incentive, from which exclusion is possible. Thus the good provided has a private and a public component and by *changing the utility function of individuals,* this approach changes the assumption that individuals posses entirely selfish preferences. Collective action is explained by altering the utility function of individuals in multiple ways in order to make provision of the collective good more attractive.

At the Network Level

• *By capturing bandwagon effects and reaching the tipping point:* By changing the type of network and the position of individuals in it we can determine to an extent, if collective action is successful.

At the Institutional Level

• This approach attempts to identify the different elements that encourage free riding and then uses an institution or institutional policies to correct the deficiencies.

---

[45] Marwell, G., & Oliver, P. (1993). The critical mass in collective action: A micro-social theory. Cambridge, MA: Cambridge University Press.

The next three sections describe each of these approaches in greater detail.

3.3. Engineering Selective Incentives for Collective Action:

This category of solutions attempts to overcome the free rider problem by incorporating excludability into the public good. It attempts to do this by the provision of selective incentives or sanctions (disincentives) that, unlike the collective good, can be applied to individual members of the group to increase their commitment levels.

However, to construct meaningful incentives for the participants in a knowledge network, it is important that we understand the mechanisms that drive individuals to participate in the first place. A large part of the reasons that motivate users to participate in knowledge networks are the same as the ones that motivate people to contribute to enterprise Knowledge Management Systems and Open Source and P2P networks. The next section provides an overview of the current research on the motivations to knowledge sharing in organizational settings is first presented. The subsequent sections in this chapter look at the motivations of the individuals participating in the Open Source Movement and the Peer-to-Peer communities and the incentive mechanisms built into such systems.

3.3.1. Motivations for knowledge sharing in Organizational Knowledge Management Systems:

The Organizational Behavior literature has recognized that the motivation and commitment of knowledge workers as a critical success factor for the implementation of corporate knowledge management systems (KMS). This is largely a result of research[46] and industry surveys[47] in the early 1990s which observed that unsuccessful KM projects have largely been those that failed to motivate their users to create, share and contribute to knowledge repositories. These studies have tried to develop theoretical understanding of motivation and commitment in the context of KMS use and implementations.

The early literature in this field was focused on the reasons for failing to meet the performance expectations of KMS. *KMS are information systems designed specifically to facilitate the sharing*

---

[46] Davenport T.H, De Long, Beers M.C: Succesful Knowledge Management Projects, Sloan Management Review, 39,2,
[47] Dyer G, Mcdonaugh B. The state of KM, Knowledge Management, v4, no5, pp 23- 41.

and integration of knowledge.[48] However, field studies indicate that employees frequently resist sharing their knowledge with the rest of the organization. In addition, knowledge is 'sticky' and does not flow easily throughout the organization even when knowledge is made available[49].The critical issue then, is understanding the social, cultural, and technical attributes of KMS that encourage and are barriers to knowledge exchange. Some of the reasons cited as barriers include insufficient communication, failure to integrate KMS into every day practice, a sense of very little personal benefit to the user, lack of time, failure to use the knowledge effectively and the difficulties of capturing tacit knowledge[50].

Malhotra and Galleta[51] provide a framework for understanding how knowledge workers' motivation and commitment affect the use of KMS and resulting organizational performance. Their paper marks a shift from the existing literature that had till then viewed commitment as a binary variable (you either cooperate or you don't), and their view tends to view commitment as a continuum from negligible to absolute commitment and from avoidance to skilled, enthusiastic and committed use by the employee. This an extension of Kelman's[52] theory of social influence that explains how social processes influence behavior. Kelman makes the point that the users moves along a continuum of processes- *compliance* (to gain a reward), *identification* (to establish or maintain relationships) and *internalization* (where the behavior is congruent with the value system).

The quality of commitment varies along the Kelman scale and in a sense Maslow's[53] hierarchy. For example the users who contribute to KMS out of compliance factors do so only because they want to maximize their incentives (and not to maximize the value added by their own contributions). In the context of knowledge networks, this would probably lead to a high volume of contributions but with low quality.

Users whose behavior is explained through the process of identification are committed because of a need for acceptance by their peers and a need for esteem within the organization. They seek

---

[48] Alavi, M & Leidner, D. Knowledge Management Systems: Issues, Challenges and Benefits. Communications of the Association for Information Systems, Vol. 1, Paper 5. USA, 1999.
[49] Szulanski, G., 1996. Exploring internal stickiness: Impediments to the transfer of best practice within the firm. Strategic Management Journal 17, pp 27-44.
[50] Alavi M and Leidner, D. Review: Knowledge Management and Knowledge Management Systems- conceptual foundations and research issues. MIS Quarterly, v25,1,2001
[51] Yogesh Malhotra, Dennis F. Galletta Role of Commitment and Motivation in Knowledge Management Systems Implementation: Theory, Conceptualization and antecedents of success.36th Annual Hawaii International Conference on System Sciences (HICSS'03) -January 06 - 09, 2003
[52] Kelman H.C: Compliance, Identification and Internalization- three processes of attitudinal change. Journal of conflict resolution, 2, 1958
[53] Abraham Maslow, Motivation and Personality, 2nd ed., Harper & Row, 1970.

to emulate the behavior of well known/powerful/popular figures within the organization. It is not yet clear if the result of the commitment brought about through identification is any different from the results of the commitment through compliance.

Commitment by internalization is the preferred way of cultivating commitment behavior. This is because a synchronization of values are more long lasting than rewards, punishment or social recognition and ensure that the knowledge worker is invested at the deepest level in the prescribed behavior. In the context of knowledge networks, the knowledge worker will try to maximize one's value added contributions and will not artificially inflate the quantity of contributions or skimp on the quality.

Another facet of internalization is that it is non-manipulative and self-governing. Both compliance and identification seek to manipulate behavior through rewards, social esteem and punishments.

Enterprise Knowledge Management Systems have historically not had tried to design explicit incentives for contribution. They have tended to rely on the inherent intrinsic motivations of the knowledge worker who was expected to view contributions as matter of pride and identity. However, recent organizational literature[54] has recognized that this view is overly idealistic and is now focused on issues dealing with the resolution of the dilemma of rewarding 'voluntary behavior'.

Deci and Ryan[55] provide a framework for understanding Malhotra and Galletta's theory on the continuum of motivation. Their work is based on self-determination theory (SDT)- a macro-theory of human motivation concerned with the development and functioning of personality within social contexts. *The theory focuses on the degree to which human behaviors are volitional or self-determined - that is, the degree to which people endorse their actions at the highest level of reflection and engage in the actions with a full sense of choice.* [56]

They argue that the traditional discourse on knowledge management which viewed intrinsic and extrinsic motivations as opposites is wrong. Instead, motivation should be viewed as *'a gradient of a knowledge worker's perceived locus of causality* (PLOC) *of specific behavior'*

---

[54] Wenger E., McDermott R., Snyder, W.M Cultivating Communities of Practice: A guide to Managing Knowledge. Harvard Business School Press 2002.

[55] Ryan, R. M., & Deci, E. L. (1991). A motivational approach to self integration in personality. Nebraska Symposium of Motivation. University of Nebraska Press

[56] Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. American Psychologist, 55, pp 68-78.

They view the regulation of behavior as being either *self-determined, controlled or amotivated-* as depicted in the figure. Though both self-determined and controlled behaviors are intentional, self-determined behaviors involve a true sense of choice while controlled behaviors are compelled by an internal or external force. One feels one *has* to do them. Amotivated actions are not intentional. Thus actions are categorized along a perceived locus of causality continuum anchored by self-determination and external control.

Type of Regulation	Amotivation	Extrinsic Motivation				Intrinsic Motivation
	Non-Regulation	External Regulation	Introjected Regulation	Identified Regulation	Integrated Regulation	Intrinsic Regulation
		Perceived Locus of Causality				
	Impersonal	External	Somewhat External	Somewhat Internal	Internal	Internal
Quality of Behavior	Nonself-Determined					Self-Determined

Intrinsic and Extrinsic Motivation as a Continuum of PLOC

Figure 2:Continuum of motivations: The Ryan and Deci framework[57]

The SDT taxonomy includes external, introjected, identified and integrated resolutions and these are all different forms of extrinsic motivations. The external regulation leads to behaviors that are performed only to satisfy an external demand or reward contingency. Introjection involves taking a regulation but not fully accepting it as one's own. Introjected regulation leads to behavior that leads to actions performed to relieve guilt or anxiety or to attain ego enhancement or pride, while Identification results in accepting or owning the action as personally important- based on self valued goals and issues. Integration occurs when identified regulations are fully assimilated into the self and made congruent with one's values and needs.

The point the above figure tries to make is that extrinsic is not the same as external(being outside the individual). Thus introjection, identification and integration – though being extrinsic regulatory styles are still internal to the individual and the consequences of feeling controlled are the same whether they are internal or external. This also partially explains how intrinsic motivations can contribute to the success of virtual communities when there is not too much extrinsic regulation.

[57] Ryan, R. M., & Deci, E. L. (1991). A motivational approach to self integration in personality. Nebraska Symposium of Motivation. University of Nebraska Press

Intrinsic motivations ensure that the processes are primarily driven by key participants for their own interest and enjoyment to extend one's own capacities and to explore and learn.

One also has to recognize that individual motivations evolve, and that much research needs to be done on why people posses such preferences and motivations in the first place. In other words, how did social preferences and individual motivations evolve? In knowledge networks, where social altruism is proposed as an important variable for collective good provision, we need to know how and specially under what conditions altruists survive free riding. Without this additional understanding the current motivation theoretical framework will provide only an incomplete answer.

The above research has important implications for understanding the motivations and incentives for participants in e-knowledge networks on development. They illustrate that it is important to understand that different motivations are important variables for the explanation of collective action and one should identify which motivations play the more important roles. Since intrinsic motivations cannot not easily observed it is hard to obtain empirical data to do this and it is difficult to come up with utility functions that support collective action. However, as we shall in our case study on GSSD, soft incentives and intrinsic motivations including commitment through internalization play a major part in determining the participation levels within a knowledge network and recognizing this could provide possible solutions to devising selective incentives for participants. See section 5.3.3 for more details.

3.3.2.   Incentive mechanisms in corporate knowledge management systems:

The incentive mechanisms in organizational KMS are usually classified as explicit (hard) or soft rewards[58].

Hard rewards offer tangible benefits to reward individuals who engage in knowledge sharing such as salary, stock options bonuses etc. Another form of a hard reward is access to information and knowledge resources within the organization. People are allowed access to information and knowledge shared by other contributors and the expectation is that you will get valuable knowledge in return for contribution to the knowledge pool. Individuals contribute to gain privileged access to this source of information.

---

[58] Hazel Hall, "Common knowledge: an exploration of social learning in distributed organizations" Presented at Managing knowledge: conversations and critiques, University of Leicester Management Centre, 10-11 April 2001

Yet another form of hard rewards are contributions in the context of career advancement within the organization. Here the focus is on discouraging knowledge hoarding and the extent to which individuals help other members of the organization. Measures indicating contributions to KMS are increasingly finding their way into performance appraisal systems.

Soft rewards on the other hand are subtler in their way or rewarding individual contributions. The focus is more on moral obligations than self-interest, and on reputation and status as being the drivers for contributions. Since individuals contribute with the aim of increasing their reputation or status, which in turn leads to career advancement, soft rewards can viewed as a more subtle case of the hard career advancement initiatives we saw earlier. Other soft rewards include a focus on personal satisfaction, altruism and pro-social behavior.

The discussion of rewards has to be framed in the larger context of the organizational factors that enable these rewards. Hall identifies environments that are most conducive to increased participation in KMS. He advises organizations to-
❑   make knowledge sharing as an explicit responsibility
❑   encourage experimentation
❑   value all contributions, regardless of the originator's status
❑   promote communities for knowledge sharing
❑   furnish employees with appropriate information and communication technology (ICT) tools

Thus firms have to use a combination of reward systems and organizational factors towards organizational knowledge sharing and learning. The question is of what to use, and when. The answer to this question is currently not clear, and we need to research companies who have tried either or both approaches. It would seem that hard (financial) rewards are useful for achieving short-term results, but soft rewards through peer recognition and increased social capital through community building would pay off in the long run. In the context of knowledge networks on development one is often faced with this choice on which way to adopt. As will be seen later, in the case of GSSD and to a lesser extent the Development Gateway, the approach so far has been to rely entirely on soft rewards as incentives for participation.

We next look for alternative answers to our questions on motivations by looking at open source projects. Open source projects are a class of knowledge networks that have innovation as their primary goal (see section 2.3- Configurations of Knowledge Networks). The open source movement is a significant economic and social phenomenon that has successfully managed to elicit thousands of top-notch programmers to contribute freely to a public good. The question is

why? What drives these people? What benefits outweigh the cost and time a developer spends in producing software? And What can we learn about the way these projects are organized and managed and apply to our discussion on knowledge networks in development?

3.3.3. Motivations and incentives in the Open Source Movement:

The open source movement originated in the 1960s at MIT's Artificial Intelligence Laboratory and the practice quickly spread through the rest of the software development community with the advent of the Internet. The rise of the movement coincides with the increasing commercialization of the software industry and the rise of monolithic software development firms such as Microsoft, IBM and Sun. The movement traces it's beginning with the discontent software developers faced due to the lack of access and the restrictions in modifying proprietary developed software.

Richard Stallman pioneered the concept of a free, open license that allows an author to grant legal rights to the users of his software to study, use, modify and distribute the software without any restrictions. This license, called the General Public License, is the bedrock of the open source movement and the GPL and its variants govern the development and distribution of most open source software products.

Open source is a term to describe the tradition of *open standards, shared source code, and collaborative development*[59]. There are thousands of open source projects in existence, with the most well known being the Linux Kernel, Perl, Apache and Freenet projects. The participation in such projects is estimated to be around 600,000 with users including individuals, firms and governments. The evolution of the open source movement is interesting because both transaction cost economics and coordination theory suggest that a proprietary software development model would be the most efficient, as it would be managed by a firm that has the capacity to ensure continued collaborative effort and minimize the impact of free riding and opportunism as well as reaping the benefits of organizational learning and other scale effects.

Open source projects fit into the category of public goods as they are both non excludable –as they are publicly accessible, and are non rivalrous- one person's use of the product does not diminish any other user's utility. However, because it is a public good, open source projects are also subject to free riding and the collective action dilemma we saw earlier. We shall attempt to understand the nature of the participants in this process, the motivations that drive them, and the structural incentives these projects provide to provide an answer to this dilemma.

---

[59] O'Reilly T. (1999) Lessons from open-source software development, Communications of the ACM, 42(4), pp 32-37

3.3.4. Norms of participants in open source projects:

The programmers who contribute to open source projects are a diverse and an eclectic group and constitute a closed idiosyncratic community. They include students, researchers, hobbyists and professional software developers. One common overarching trait in this group is that they all prize autonomy, independence and self-determination.

They are also referred to as hackers- a term Eric S Raymond defines as those who love programming for the sake of doing it, for the sake of obsessively solving a problem. Highlighting the difference between hackers and traditional software developers, Raymond [60] argues that *"The 'utility function' Linux hackers are maximizing is not classically economic, but is the intangible [product] of their own ego satisfaction and reputation among other hackers."*

Hackers have been portrayed both in the news and in fiction as unique for their cultural proclivity to reject the materialist trappings of modern life[61]. This view, popularized by Steven Levy in his book Hackers: Heroes of the Computer Revolution, and the further romanticization of hackers by Neal Stephenson in his books -Crypronomicon and Snow crash, that portrays hackers as the elite in a society where social boundaries are drawn in terms of technical expertise is further reflected in the fact that engineering students place great value on things - books, movies, ideas - that connect their concerns with something larger.

Although generalizing about the norms within this diverse group can be misleading, there are certain common norms that govern this community-a sense of belonging to an elite group, an obsession with freedom of speech, a feeling that respect must be earned and cannot be derived from position and a distaste for bureaucracy and administration.

These set of norms affect the process of software development in different ways. In Gerald Weinberg's classic The Psychology of Computer Programming looks at the issue of 'egoless programming' in projects where developers are not territorial about their code- a direct result of the feeling that quality of the works is what matters, and that position and status have nothing to do with the project. This encourages other people to look for bugs and potential improvements in the project, resulting in improvements that happen faster than in professional software development organizations. On the other hand, a distaste for administrative activities leads to the problem of setting up formal/informal governance procedures for governing the process of development and its documentation.

---

[60] Eric S. Raymond. The Cathedral and The Bazaar. In The Cathedral and The Bazaar, pp 27--78. O'Reilly and Associates, Sebastopol, 1st edition, 1999
[61] David Lancashire: Code, Culture and Cash: The Fading Altruism of Open Source development, pp 2-3 First Monday, volume 6, number 12 (December 2001),

To better understand this resistance, one has to look at who really owns the software, as ultimately administrative and governance procedures will have to flow from him. In open source projects, the owner is the person who has the right to *distribute* modified versions of the software. The ways to acquire ownership are by either founding the project, or when the ownership is handed to a person by the founder or by assuming ownership when there is work to be done, and the previous owner has lost interest. None of these are hard rules, but the norms of the community recognize ownership only through direct involvement with the origin(or the rebirth) of the project.

Since ownership of a project also carries it with connotations of technical expertise, experience and interest, the community largely follows the governance procedures set forth by the owner. However, a few high profile projects have been discontinued or are faltering because the participants perceive excessive control and interference in their contributions. The issue of ownership also brings us to the question of leadership styles within the open source community. The most famous leader is of course Linus Torvalds, the poster boy of the open source movement and the founder of Linux. Leadership in open source projects cannot be coercive and it essentially involves nurturing interest in the project till it becomes self-sustaining by recruiting and energizing participants. As Raymond[62] puts it *"the cutting edge of open-source software will belong to people who start from individual vision and brilliance, then amplify it through the effective construction of voluntary communities of interest".*

3.3.5.   Motivations for contributions to Open Source Projects:

Borrowing from our earlier framework used to characterize motivation in organizational settings, one could divide the motivating and regulating factors in to intrinsic and extrinsic components.

Deci and Ryan posit that the theory of **intrinsic motivation** derives from a human need for competence and self-determination which in turn is directly linked to the emotions of interest and enjoyment. In January 2002, Lakhani and Wolf[63] using a web based survey, administered to 684 software developers in 287 F/OSS projects, have found that intrinsic motivations are the most important drivers to contribution. Some of the results of this survey are reproduced below-

---

[62] Eric S. Raymond. The Cathedral and The Bazaar. In The Cathedral and The Bazaar, pp 5-7. O'Reilly and Associates, Sebastopol, 1st edition, 1999

[63] Lakhani, Karim and Wolf, Robert G., "Why Hackers Do What They Do: Understanding Motivation and Effort in Free/Open Source Software Projects" pp 11-19 (September 2003). MIT Sloan Working Paper No. 4425-03.

Motivation	% of respondents indicating up to 3 statements that best reflect their reasons to contribute (%)	% volunteer contributors	% paid contributor	Significant difference (t statistic/p value)
Enjoyment based Intrinsic Motivation				
Code for project is intellectually stimulating to write	44.9	46.1	43.1	n.s
Like working with this development team	20.3	21.5	18.5	n.s.
Economic/Extrinsic based Motivations				
Improve programming skills	41.3	45.8	33.2	3.56 (p 0.0004)
Code needed for user need (work and/or non-work)*	58.7	-	-	-
- Work need only	33.8	19.3	55.7	10.53 (p 0.0000)
- Non-work need	29.7	37.0	18.9	5.16 (p 0.0000)
Enhance professional status	17.5	13.9	22.8	3.01 (p 0.0000)
Obligation/Community based Intrinsic Motivations				
Believe that source code should be open	33.1	34.8	30.6	n.s
Feel personal obligation to contribute because use F/OSS	28.6	29.6	26.9	n.s.
Dislike proprietary software and want to defeat them	11.3	11.5	11.1	n.s
Enhance reputation in F/OSS community	11.0	12.0	9.5	n.s.

Table 3: Summary of the Karim-Lakhani survey results[64]

Lakhani and Wolf divide intrinsic motivations into enjoyment-based motivations and community based motivations. **Enjoyment based intrinsic motivations** were first studied by Csikszentmihalyi, who put forward the theory that some activities are pursued for the sake of enjoyment derived from doing them and enjoyable activities are found to provide feelings of "creative discovery, a challenge overcome and a difficulty resolved". In their survey, Lakhani and Wolf find that a large proportion of open source participants (over 45%) find that participation in open source projects is intellectually stimulating.

There is also strong evidence of **community based intrinsic motivations** in open source projects where there is a well-defined community goal, and the participants exhibit strong collective identities. There is a shared sense of meaning and collective identity centered on the concept of a hacker culture, and the motivation of participants is to act consistent with the perceived norms of the group. Along with the feeling of belonging to this group, there is also a desire for reputation and a need to be acknowledged by the rest of the community. Together, as the survey indicates- the belief that code should be open, the desire to beat proprietary software, a stake in the reputation of the open source movement, and a feeling of belonging and being recognized by the community, constitute they key drivers that motivate this community.

Although the popular impression of open source projects is that the contributors are solely driven for the fun of it, there is a strong **extrinsic motivation** component for participation too. Some of the reasons for participation may include developers who are paid to participate, a user who feels

[64] The Boston Consulting Group Hacker Survey Release 0.3: The Boston Consulting Group, January 31, 2002 – Distributed under the GNU Free Document License v1.1

the need for a particular software or a significant professional interest in how a software should be developed, what functionality it has etc., and even factors such as career advancement (where the status of being involved in an open source project might lead to signaling in the job market of their software development skills) and increased learning. However, as the above figure indicates, the intrinsic motivations that drive the volunteers to contribute are also prevalent among the paid to contribute members.

In fact, the Lakhani-Wolf survey indicates that learning is the most important extrinsic motivation among participants (both paid and unpaid), and there is a feeling that participating in an intense peer reviewed process would greatly improve their programming skill. This is different from the viewpoint of learning as a source of intrinsic motivation –where the inherent creativity in the process of knowledge construction and learning can serve as a source for enjoyment based intrinsic motivation.

Continuing with the use of our framework for characterizing motivations and incentives in corporate KMS, we now look at the **organizational factors** that influence and enable the rewards expected by the participants.

One of the preconditions for the creation of open source organizations is a large community of practice with a strong, shared culture. Given the nature of the participants the open source movement has attracted in the past, this is not a difficult thing to provide. However, recognizing the nature of the participants and their norms and preferences, the earlier formal structure, governance and processes that were traditionally used to manage software projects will prove to be ineffectual in this environment. There need to be mechanisms for workers to complete tasks through informal relationships and networking.

For example, the **governance** of this community must be managed in a way that is perceived as being fair and equitable by the participants. This entails a transparent governance mechanism, relinquishing central command and control structures and allowances for working in teams and to make decisions by discussion and voting. Agenda for the community should be set both top down (when the project is initially formed) and bottom up –where members can choose what they want to do based on their skills, competence and knowledge.

The **infrastructure** for software development and project management, product release and documentation should be consistent with the nature of interest amongst participants and their level of motivation. For example extensive documentation demands can diminish motivation and participation. Knowledge management tools like CVS-a central repository in which the artifacts

are stored and managed and protocols for adding and retrieving artifacts from the repository, collaboration mechanisms for peer review and discussion need to be in place.

The nature and the **process of recruitment** for the project should be carefully managed. Most open source projects allow for new members to join and current members to leave the community freely. This flux within the project would have to be managed carefully. Recruitment and advancement within the project should be wholly merit based.

The project also has to set up meaningful incentives and rewards after understanding the key motivating factors of the participants. The challenge is devising a reward system where the community as a whole is responsible for its work and gets rewarded and penalized collectively. For this to work, community members will have to develop a high level of trust among each other. Reputations and designations within the project are the most important rewards currently used by the open source community. Designations such as ubercoder and linux kernel team member are viewed as prestigious positions to aspire to and even the promise of a penguin t-shirt can go a long way. There has to be a carefully planned path for progression from contributing member to active developer to core project team member.

The open source software (OSS) model is a fundamentally new and revolutionary way to develop software. Although a number of benefits can be realized by adopting the OSS model, it may not be suitable for all organizations, for all projects and for all people. An understanding of the motivations of people who contribute to the project is the first step necessary to ensure meaningful sustained contribution.

This discussion so far indicates that there are a number of tradeoffs that might face knowledge networks trying to adopt a similar model. This includes making difficult decisions on- the extent to which the community should be self governed, the extent of selectivity in recruitment, challenges involved in matching organizational processes to suit motivation and involvement levels etc.

There is much innovation-based knowledge networks (like the ThinkCycle initiative at MIT) can learn from the way OS projects are structured and managed. Knowledge Networks focusing on innovation as their end product are particularly reliant on the quality of contributions into their knowledge base. The evolution of the role definitions, processes, checks and balances within the OS model over the years can provide valuable insight into the way similar processes can be re-engineered into other knowledge networks in the sustainable development domain.

One critical enabling condition we saw from the OS model is the need for knowledge networks focusing on innovation on sustainable development issues is the need for a strong, shared culture. This culture exists in the offline world- partly because people involved in development work to an extent share the same values and norms and also partly because of the normalizing effect international development institutions have had in creating a shared vocabulary and meaning on issues. The challenge is to replicate this culture within an online community.

We next turn to another unique model of a virtual community- the peer-to-peer networks used for sharing files on the Internet. The peer to peer (P2P) network is not a conventional knowledge network as the exchange mechanisms do not involve the transfer of knowledge in any fundamental way. However, it can also be viewed as a virtual community with the same elements of collaboration and the problems of collective action and free riding we observed in our discussion on open source and corporate knowledge management systems. Free-riding in the context of P2P networks is more serious than free-riding in other models of knowledge networks-where free riding is tolerated, and in some cases even encouraged- the logic being that as the cost of knowledge provision is independent of the number of users served, free riding does not pose any direct cost. However in P2P networks, free riders by consuming valuable resources in the network (bandwidth/space) pose a direct cost on all other users within the network, thus making the problem much more serious.

3.3.6. Motivations and Incentive structures within P2P file sharing networks.

The word P2P was brought into the mainstream vocabulary virtually overnight with the emergence of the Napster phenomenon. Napster-a software designed and developed by Shawn Fanning, an 18-year-old freshman at college, was aimed at relieving the difficulties in locating and swapping digital music files online. The first version of Napster combined the instant-messaging system of IRC, the file-sharing functions of Microsoft Windows and Unix, and the advanced searching capabilities of various search engines. And the software was free.

The popularity of Napster and its cousins Gnutella, Kazaa and Limewire is indicated by the fact that 60 million Americans trade files online, and that 40% of Internet traffic is because of P2P applications. However, this huge growth of P2P and the illegal trading of copyrighted files has proved to be a source of concern for record companies -represented publicly by the RIAA (Recording Industry Association of America), and they have managed to successfully shutdown Napster and as we speak KazaaLite. The RIAA has also taken the war to the users of P2P systems and has issued 261 subpoenas against high profile file sharers. Artists are divided in their support for P2P applications, some of them feel that P2P provides a medium for fans to

sample their music before buying, while a couple of high profile artists publicly condemning it as tools for stealing music.

Since Napster, the P2P movement has moved beyond file-sharing, to such projects as SETI@home at UC Berkeley and the Stanford Protein Folding project, which pools the spare processor power of the participants to process information. Our discussion will focus on only file sharing P2P networks.

Although there has been a proliferation of companies offering P2P clients for file sharing, there have also been instances of many high profile networks who have closed shop or are stagnating because of their failure in securing enough cooperation to become truly useful. In such communities users stop contributing and only consume resources. This free riding behavior is similar to the ones we have observed before in our discussions on other kinds of virtual communities.

In the context of P2P networks, the common good is the provision of a very large library of files, music and other documents to the user community. There are other second level public goods like the shared bandwidth in the system. The dilemma for each individual is then to either contribute to the common good, or to shirk and free ride on the work of others.

Adar and Huberman[65] have also identified another problem caused by free riding. If only a few individuals contribute to the public good, these few peers effectively act as centralized servers. Users in such an environment thus become vulnerable to lawsuits, denial of service attacks, and potential loss of privacy. This is relevant in light of the fact that systems such as Gnutella, Napster, and FreeNet are depicted as a means for individuals to rally around certain community goals and to "hide" among others with the same goals. These may include providing a forum for free speech, changing copyright laws, and providing privacy to individuals. Their experiments and surveys that free riding is rampant in P2P networks - In Gnutella- an open source P2P client, a large proportion of the user population, upwards of 70%, enjoy the benefits of the system without contributing to its content.

[65] Eytan Adar and Bernardo A. Huberman. Free riding on gnutella. Technical report, Xerox PARC, 10 Aug. 2000. http://citeseer.nj.nec.com/adar00free.html

3.3.7.   Motivations of File Sharers:

Over 60 million Americans are breaking the law every time they get on to Kazaa or Gnutella. What could explain this phenomenon? It is quite easy to understand the demand side of P2P applications. People want a song, and they download. What is puzzling is why people would want to share or upload their songs? Some of the explanations for this could be –

- Intellectual Law is complex and people really don't know if they are breaking the law by sharing files(not so valid any more as recent press coverage of lawsuits, and surveys that indicate that 40% of people believe what they are doing is illegal)

- People tend to ignore the law when sometimes, when they feel that the consequences are not too serious (Shirky[66] compares this to people driving slightly above the mandated national 55 mph speed limit- which doesn't exist anymore because no one really followed it)

- It doesn't feel like theft, it feels like a swap fest where I can visit and explore interests with like-minded fans. Sharing files is a good way to know about others and others to know about me and my interests.

- Anger over being ripped off by arbitrary pricing and packaging by the music industry. Resentment towards big record companies for failing to serve eclectic tastes and over the fact that copyright law protects record labels more than the artists- who seem to end up with very little of the proceeds[67].

- Fans are cheapskates, and no one would refuse a free deal and P2P is just an excuse to avoid paying for music. Some P2P apps. require users to upload a certain number of files before they are eligible for download, and I don't mind uploading so long as I get my free downloads.

- Sheer convenience, the music industry has still to recognize the convenience P2P software provides in searching for and downloading music on the web. This is now changing with legal downloads possible through Apple's iTunes, and other offerings.

Another important reason for file sharing has been the evidence of a 'gifting culture' within P2P networks. In the context of knowledge networks, when the resource in question is plentiful- here knowledge, then the chief measure of status is the ability to gift knowledge to others. The following section explains this in greater detail.

---

[66] http://www.shirky.com/writings/napster_speech1.html
[67] Courtney Love does the math: Salon Magazine, June 14, 2000

3.3.8.   The Gift culture within P2P networks

One can characterize the community culture within file sharing networks as a gifting culture (as opposed to an exchange culture). Exchange culture is what we are familiar with: exchanging money for goods, while Gift culture occurs when anything of interest exists in abundance. The standard example given is that of a small group living in a jungle where all foods and materials known exist in abundance. In such a community, status is determined by one's ability to give away goods. Giesler and Pohlmann [68] further define gift culture as a *"fundamental human impulse to display, to share, to bestow,"* and a *"deep tendency to create social ties through exchange of gifts."*

Giesler and Pohlmann then explain how communities are constructed through the process of giving and how social ties are created through the exchange of gifts. Using this framework, they analyze the motivations for user downloading and uploading by characterizing gifting behavior as Either of (at the individual level)

Gifting as realization, where the primary benefit is functional and sharing has an individual utilitarian purpose for the user. This is more true for downloads and in some cases uploads(amateur artists putting up their work online)

Gifting as purification, here purification is a metaphor used to describe *'gifting as a form of resistance against the influence and impacts of the contemporary music entertainment regime'.*

..and at the community level..

Gifting as participation, where motivations are drawn from the impulse to belong and integrate. The file sharing community is seen as a movement, which one wishes to join in order to add value to others.

Gifting as renovation, emphasizes the P2P network as a *locus of communally enacted social change.* Gifting is attached to a political matter of concern and is used to build up and communicate an alternative to the regime.

All of the above reasons are true to some extent and explain user motivations in Knowledge Networks. The reasons given above fit broadly into the categories of people who are motivated by

---

[68] Markus Giesler, Mali Pohlmann: The Anthropology of file sharing: consuming Napster as a gift, pp 23-27 Submitted as a Competitive Paper to be published in the ACR 2002 proceedings

commitment to a cause, materialistically motivated, altruistically motivated, motivated by glory or motivated for utilitarian purposes.

To better understand these motivations, it would be useful to learn more about the demographics of the file sharing user population. According to a Times/CBS News poll- a significant proportion of file sharers are young (55% below 28 years) and are enrolled in college (43%). However, the trend has spread across nearly every demographic group, more and more people in the 30-50 age group getting involved. Another more startling study released in July by the Pew Internet and American Life Project shows that among the 35 million adults that its survey indicated download music, 23 million said they did not much care about the copyright on the files they copied onto their computers. Among the 26 million who made files available for others to copy, 17 million did not care much about whether they were copyrighted. It is not clear on what section of the populations is motivated by what factors. One would expect the youth and college students to be motivated because of utilitarian features and to a lesser extent commitment to the cause of sharing free music and community building around their favorite artists. However, research still needs to be done to validate this.

In the case of knowledge networks on development, the demographics of file sharing is expected to be different- a demographic analysis of participants in the Development Gateway topic pages is provided in Section 4.5.5. The most useful contributors in terms of knowledge value added are likely to be those who are experienced in a domain within sustainable development and who are higher up in their organizations. Although figures are not available on membership by 'age group', it is expected that the median age of members is to be much higher than that of P2P networks. Also, if one were to divide participation into two categories – contributions to a knowledge pool (including articles, comments on articles, and discussions) and participation in more utilitarian activities –like posting/replying to jobs, posting announcements etc. it is to be expected that older experienced members are more likely to participate more because of their commitment to the cause of the network and to the knowledge pool than due to its utilitarian features. However, this hypothesis is yet to be validated.

3.3.9.   Summing up:

Looking back at the motivations of user contributions in the three very different virtual communities we have discussed so far, we see that the motivations for sharing in corporate KMS, open source software development and in P2P networks fall broadly under the categories of altruism, group commitment, reciprocity, reputation, self-efficacy and utilitarian factors.

Practitioners who are in charge of designing and managing knowledge networks need to understand what really drives their users and devise incentive mechanisms to recognize these. Designers should recognize that they could exercise substantial leverage by enabling organizational factors that favor these motivations. For example, one of the recommendations for GSSD as explained in section 5.3.1 is to create a greater sense of community within the network. Based on our discussions in the previous sections this would involve creating capabilities for future interaction and a record of past interactions, identity persistence – to build trust in the community; create well defined and defended group boundaries – to increase the granularity of a community and create mechanisms for increased visibility and recognition of contributions.

So far we have looked at individuals and their motivations towards knowledge sharing and contribution. In the next section we shall look at the knowledge sharing process at the network level, where we discuss how structural solutions could be brought about to solve the collective action problem in knowledge networks.

3.4. Structural Solutions to the Collective Action Problem:

Structural solutions depart from the solutions that tackle collective action problems at the individual level(and the focus on the norms, behaviors, motivations and incentives), by focusing on the mechanisms by which the preferences of individuals interact and aggregate in the network. Originally proposed by Granovetter (1978)[69], this kind of solutions deal with the issue of group composition (group size and the distribution of interests and resources etc.) and intra-group interactions. Olson differentiates between three types of groups: *privileged, intermediate*, and *latent*[70]. A privileged group can be thought of as a group in which one of the members receives a disproportionately high utility from the consumption of the collective good, and he would be willing to produce a big share of the good by himself. In such cases, the collective action problem would be solved without any need of re-organization or restructuring.

The intermediate group lacks such a group member; however, it is still possible for a small number of individuals to coordinate their actions and produce an almost optimal amount of the good. The key to the existence of an intermediate group is that because the group is not big enough, the actions of an individual can considerably affect the utility of the other members and structural solutions might be enough to facilitate an adequate provision of the collective good.

Latent groups on the other hand are too big for their members to have a considerable effect on each other's utility. This results in the free riding problem and the provision of the collective good is doomed to fail. This also corroborates the empirical evidence that shows that group Size has an impact on the outcome of collective action problems[71].

However, this approach of large groups always fail while small groups succeed is not a wholly correct one. One also has to take into account the group composition, both in terms of the distribution of interests amongst its members as well as the distribution of resources. Interest heterogeneity leads to what Olson calls the "exploitation of the great by the small". Hardin[72] also argues that collective action is more likely in groups where interests are highly heterogeneous. Resource heterogeneity has received lesser attention in the literature, and Hardin argues that political processes tend to produce goods that mainly benefit wealthier people.

---

[69] Granovetter M. (1978) Threshold models of collective behavior. American Journal of Sociology. 83(6), 1420-1443.
[70] Olson, M. (1965). The logic of collective action. Cambridge: Harvard University Press.
[71] Spilerman(1970): "The causes of racial disturbances: A comparison of Alternative explanations" American Sociological Review 35: pp 27-49
[72] Hardin, Russell. 1982. Collective Action. Baltimore: The Johns Hopkins University Press.

The implications of group composition and heterogeneity is that although on an *average* group interest and resources might be low, it is still possible for a heterogeneous group to achieve collective action if the *form of the distribution* of resources/interests is right. Thus, greater variance with a positive skew might explain the prevalence of collective actions in situations like political lobbying, where the costs might be prohibitive for an average individual to take any action.

One way to characterize knowledge networks would be by a set of interdependent decisions made by their participants(with diverse interests and resources) who take into account how much others have contributed to the knowledge pool and then make up their mind about contributing to the collective action.

Thus each individual would be faced with the question of 'how much return can I expect with a certain degree of contribution?'. More interestingly, it would be useful to know an answer to the question of 'If a person contributes twice the level of resource/time to a knowledge network , does this raise the probability of the success of the community/network by two times/less than two times/more than two times?'

Economists typically model the response to such a question in terms of 'production functions'. Some common found production functions are-
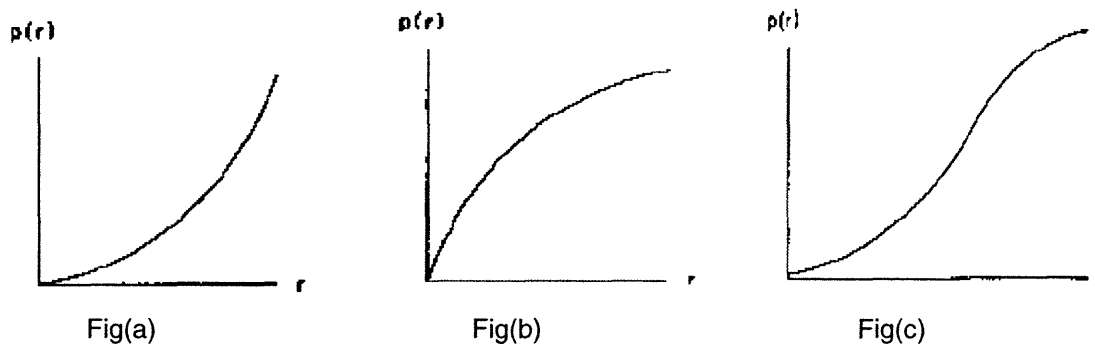
Figure 3: Types of Production functions

Figure (a) represents increasing returns to scale, where the marginal value of network grows with each additional resource – successive contributions generate progressively larger payoffs, while Figure(b) represents decreasing returns to scale- where the marginal value of the network decreases with growth in content – these are typically found in networks with declining participation.

To understand production functions in knowledge networks, imagine the incentives to contribute in the early stages of community formation. A volunteer contributing an abstract into GSSD or the DG at an early stage is less likely to benefit from the small database. It however gets easier to entice voluntary contributions once the database grows in size and thus, the production function for knowledge networks would mostly be accelerating in nature in the beginning. The value of a community increases with the number and the extent of topics covered and the quality of contributions.

We can see that adding one, two or five resources in the early stages does not increase community's perceived value steeply. However, when there are a few thousand contributions, a large population would find the network valuable. This would translate to increased participation (at least in terms of downloads and visits). While an individual's contribution would add little to the production function, the number of individuals contributing towards the collective good would enhance significantly at this stage resulting in the production function taking off in an accelerated manner. At this stage, the network is expected to have reached critical mass – defined here as the size of the network where it begins to 'self-catalyze'. The challenge of critical mass occurs when to be regarded as of high value by a wider audience, the community requires a critical mass of contributions and users. While it is a challenge to predict the critical mass, understanding the mechanisms by which it occurs give us an insight into how we can make knowledge networks sustainable.

3.4.1. The theory of Critical Mass:

The theory of critical mass has been used to explain models of collective behavior in situations that describe herd behavior in mobs, and the mechanisms by which standing ovations 'spread' in theatres. This theory also borrows upon similar work on the contagion effect in infectious diseases, the spread of rumors, the extent of participation of workers in strikes, voting behavior etc.

In his seminal work on Threshold models in Collective Behavior, Granovetter[73] provides the example of choices individuals face in a riot situation, where he explains that the costs and benefits for an actor to participate in a riot depends in part on how may other people participate in a riot. As the number of people joining a riot increases, the costs for an individual decreases (as the probability of being apprehended is smaller in larger crowds).

---

[73] Granovetter, Mark. "Threshold Models of Collective Behavior." American Journal of Sociology, 1978, 83(6), pp. 1420-1443.

He argues that different individuals have different thresholds for action. A person's threshold is reached when the benefits of joining a riot are greater than the costs of participation. Thresholds are different from norms but result from in part from them. Thresholds will arise from the same causal variables that impact behavior (background, education, occupation, social position etc.), but the important distinction is that thresholds are also situation specific. Since thresholds are computed from the costs and benefits depending on the situation, and different configurations can lead to different actions.

Operationally, in the critical mass theory, a person's threshold is defined as the proportion of the group that would have to participate before he joins the riot. Hence a 'radical' would have a low threshold (and an extreme radical would have a threshold of 0%- he will riot even if no one else does), conservatives would have a high threshold (ultraconservatives- individuals who wouldn't participate even if every one else participates would have a threshold of 100%), and individuals are distributed along this continuum of thresholds.

He illustrates the concept of critical mass by a simple example where individual thresholds are uniformly distributed. Assume 100 people are milling round a town square, with their thresholds distributed as follows- there is one individual with a threshold of 1, one with a threshold of 2, and so on, till the last individual has a threshold of 99. The person with a threshold of 0, will instigate the riot, this in turn will activate the person with a threshold of 1 in to action, and so on, until all 100 people have joined. If we were to change this distribution into one where the individual with a threshold of 1, is replaced by another with a threshold of 2, then a riot does not have a mechanism to proceed.

The implications of this result is that one cannot extrapolate the aggregate outcome from individual dispositions. Granovetter explains this nicely when he says that the press if they were to write a story the next day on the two outcomes would describe the first situation as "A crowd of radicals engaged in riotous behavior', while the second situation would be described as 'A demented troublemaker broke a window, while a crowd of solid citizens looked on'.

Thus threshold behavior could be useful in situations where the average level of collective action could indicate a certain outcome, but the action is usually not seen. In the context of knowledge networks, this result implies that it is not enough for every one in the group to have a sufficiently enough aggregate feeling for contribution, but the distribution of thresholds for action would also have to be considered.

Granovetter also gives a mathematical formulation for the phenomenon described above. If we denote thresholds as x, and the frequency distribution as f(x), and the cumulative distribution function (c.d.f) as F(x) – where the c.d.f is the proportion of the population with a threshold less than or equal to x, and the proportion of a population who have joined a riot by a time 't' as r(t), then it is possible to find out if the equilibrium point can be reached. The results of his exercise for a normal distribution of thresholds is reproduced below.
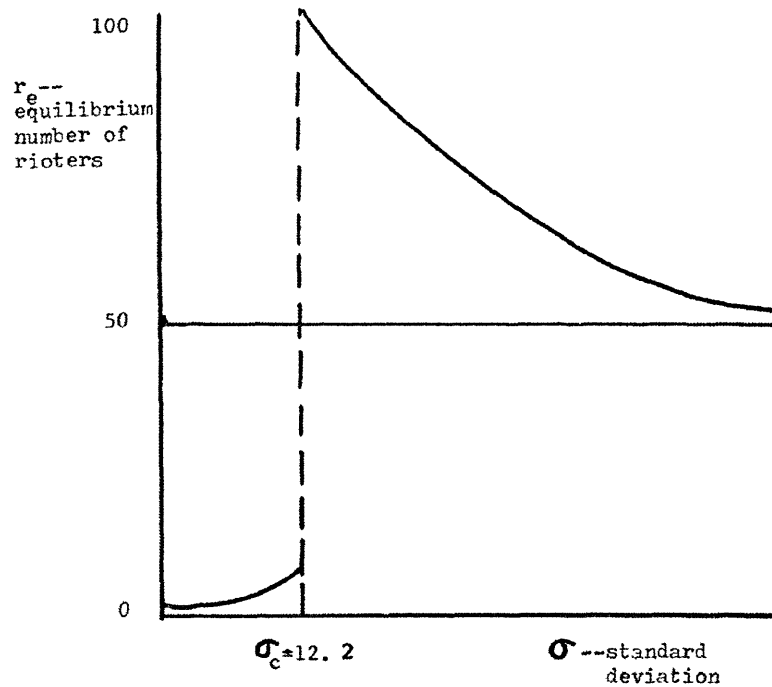
Figure 4: Instability of the critical mass equilibrium[74]

The figure above indicates that even a slight perturbation about the critical standard deviation leads to a huge change in the level of equilibrium. Thus even minor changes in the composition of a group and their thresholds could affect their level of collective action, and two crowds with the same average values of costs and benefits could still vary wildly in their outcomes.

3.4.2.   The nature and determinants of thresholds:

The distribution of preferences and thresholds and the resultant equilibrium point is in turn dependent on -

- The 'social structure' of the people in the network

    The influence of any given person has on another's behavior could also be influenced by the nature of the relationship and the social structure within the group. Factors that need to be considered are the influence of trust, friendship and influence. Questions that are still to be

[74] Reproduced from Granovetter M. (1978) Threshold models of collective behavior. American Journal of Sociology. 83(6), 1420-1443.

answered include - Does a high or low density of friendship ties have a greater effect in modifying the equilibrium result for some particular distribution of thresholds?, What is the impact of strong ties, symmetricity amongst ties etc., in influencing the equilibrium. It is yet to be determined if the strength of social ties have a positive or negative effect on the equilibrium outcome, but the overall goals of such an analysis is to specify the impact of social structure on collective outcomes.

and

- The ways in which the actors are distributed spatially and temporally

Granovetter's model assumes that each individual is completely connected to all other actors of the network —regardless of the spatial and temporal dispersion of the aggregation. This is incorrect. In riots, spatial limitations could influence the extent to which a rioter's actions are viewed by and evaluated by other participants in the group. Similarly, in knowledge networks, people rarely act immediately after seeing the behavior of another actor. The actions are more akin to all actors acting at once on a particular level of contribution, than on a sequence of interdependent decisions. Modeling the influence of spatial and temporal distribution in threshold models is complex, and presents mathematical challenges.

Also, the ways in which people enter and leave the network also affect the equilibrium outcome. This might have implications to the stability of the network and lead to unstable equilibria when the standard deviation of preferences vary and come near the critical standard deviation point.

3.4.3. Implications for re-engineering knowledge networks

The theory of critical mass provides us with an important insight that not every one needs to be mobilized to action since free-riders do not necessarily reduce the benefits to those who volunteer. If one were to go by the Prisoners Dilemma payoffs we saw before, there would be no hope for collective action. However, critical mass theory says that it is not necessary to synchronize the actions of everyone in the group, only enough members to create a critical mass would suffice.

This has implications for the allocation of resources, recruitment of members in to the knowledge network, and on the solidarity in large groups.

On **resource allocation-** consider the context of the Development Gateway (DG) at the World Bank. The DG has over 35 topics for discussion on all manner of things related to sustainable development. There are over 15 'administrators- including topic guides, advisors, editors etc.' who manage the discussion and contributions of members in each of these

forums. The current level of contributions in the topics are not very encouraging. There is also a resource constraint in terms of the time and cognitive demands on the administrators who moderate these groups. Their earlier efforts in improving the level of contribution in each group-involved contributions to each of the groups they moderate. This would usually amount to 2 or three new topics or follow-ups for each of the groups they moderate. Using the critical mass framework, one would argue that their effort would be better utilized if they were to spend more time on a subset of the groups, rather than spreading their contributions over so many forums.

Again using the critical mass theory, one would assume that if they were to '**recruit**' certain key contributors in each of these categories using selective incentives, then there is a higher chance of successful and engaging discussion in these topics. This approach has been used by the MIT developed GSSD which relies on key institutional partners who provide a steady supply of content.

Critical mass theory also provides us clues on achieving group solidarity by managing the heterogeneity in a group. Critical mass theory says that the larger the size of a heterogeneous group, the larger the number of potential benefactors, and hence more the chance of achieving critical mass. By attracting a diverse user base, the chances of outliers who will participate and increase contributions overall is increased. In practice, this means attracting members from outside the target user base makes sense if they can bring about a skew in the distribution of thresholds to participation in the activity. The caveat here is that the new people that are included should contribute meaningfully. Otherwise this could put off the core user group so much that it could decrease contributions from previous levels.

Another way of attaining critical mass is to create 'crisis' (extraordinary events that affect a large portion of the group) situations that increase the level of contributions dramatically in a very short period of time. Once contributions reach a certain level, it can be sustained by the natural 'inertia' of the group.

Critical mass theory also says that start up problems explain why so many collective action problems fail. Start up problems occur because of low yields at the beginning. Marginal utility theory says that very few people will contribute to the public good at this stage and even rational actors who have a stake in the network will prefer to stay silent rather than serve as a 'loss leader' to the group. To break out of this deadlock, one needs to provide either high rewards or sanctions for contributing/not contributing at the very beginning. The key task of the organizers is to raise expectations at this stage and show the participants what the network could become. Once the level of contributions start to increase, the problem

becomes one of 'follow up', and in changing the incentives for participation continuously as the level of contribution changes.

The chief criticism of critical mass theory is that it assumes rational actors who have an idea of their thresholds, costs and benefits. This is not always the case. Also nothing is said about the instability of the equilibria attained at critical mass. Other criticisms leveled are that the spatial and temporal effects, effect of friendship and social structure all make it impossible to model using real data. However, critical mass theory still provides us with a deeper understanding of how solidarity can thrive in large groups and gives us valuable insights on managing knowledge network from the inception through growth and maturation phases.

3.5. Institutional Solutions to the collective action problem:

The institutional approach focuses its attention on the different elements that encourage free riding and then uses an institution to correct the deficiencies.

Ostrom[75] defines institutions as the set of working rules that are used to determine who is eligible to make decision in some arena, what actions are allowed or constrained, what aggregation rules will be used, what procedures must be followed, what information must or must not be provided, and what payoffs will be assigned to individuals dependent on their actions.

In their paper[76] on organizing groups for collective action, Dawes et al studied the payoffs in the prisoner's dilemma and looked at the incentives to defect. They identified two reasons for a player to defect 1. the "fear" of getting the lowest payoff by cooperating when the other player defects, and 2. due to "greed" – getting the highest payoff by defecting when the other player cooperates. They then looked at the possibility of creating an institution that controls either a player's "fear" or "greed" and looked at whether free riding might be prevented.

They also point out to the existence of institutions in everyday life which already do this. For example the use of a "money-back" policy eliminates the "fear" of receiving the lowest payoff since it returns your contribution in case the other does not. They give us the example of how this money back guarantee was used in the efforts by the faculty at the author's university in

---

[75] Ostrom, E. Governing the Commons: The Evolution of Institutions for Collective Action. Cambridge: Cambridge University Press. (1990)

[76] Dawes, R. M., J. M. Orbell, R. T. Simmons, and A. J. C. van de Kragt (1986). Organizing groups for collective action. American Political Science Review 80: pp 1171-1185.

hiring a lobbyist. The desired lobbyist- a public good as the results of the lobbying would flow to everyone in the university regardless of whether they contributed to hiring him or not, required 30,000 dollars for his work. The university asked all the faculty members to chip in, with the explicit promise that the money would be returned if less than the $30,000 was raised. The solicitation was successful.

They also refer to "enforced contribution" as a mechanism that eliminates a player's "greed" for the highest payoff because it imposes a fine such that defecting while others cooperate does not give her a higher payoff than cooperating. They provide us an example of how apartment dwellers resist developers who wish to convert their apartment building into a condominium. The developers offer to sell the units at a reduced price to anyone who wishing to vote for the conversion before a particular deadline. One contributes to the apartment dwellers efforts by withstanding the offer. If the efforts fail, and the conversion proceeds, those who withstood the offer are out of pocket to the extent of the reduced offer. However, it is not possible for someone to free ride on the restraint of others because, if a certain number withstand the offer, the conversion will not occur and he will not stand to gain.

Other authors have done similar analysis. In some cases, to investigate the effects of existing institutions such as third party monitoring and suggesting new institutional settings and centralized regulations that might successfully promote collective action. However, Ostrom argues that asserting that central regulation is necessary tells us nothing about the way a central agency should be constituted, what authority it should have, how the limits on its authority should be maintained, how it will obtain information, or how its agents should be selected, motivated to do their work, and have their performances monitored and rewarded or sanctioned.

Institutional solutions focusing on collective action in knowledge networks has focused on the roles and responsibilities of the sponsoring organization. Some of the key questions that have to be answered to effectively manage a community of knowledge sharing include
- What level of control should the sponsor firm maintain?
- What are the roles and responsibilities of the sponsor?
- What are the criteria for selecting the members?
- Can individual contributions be rewarded?
- How should the community be allowed to evolve?
- What incentives promote long-term member involvement?

There are no easy answers to these questions. Organizations have experimented with different institutional structures for knowledge networks – from the Centrally developed loosely controlled Development Gateway, to the decentralized but tightly controlled GSSD. In general the level of control by the sponsoring organization should be such that it pulls individual members and organizations into a circle of shared concerns. The level of support provided by the sponsor firm will vary with time, from initial physical support and attracting new members to maintaining high levels of commitment from the participants. The services offered by sponsor should be tiered, depending on intensity of the individual members' involvement.

3.6. Summing Up

Our discussions so far have centered on solutions at the individual, network and the institutional level in solving collective action problems within knowledge networks. The choice of which solution (or combination of solutions) to use is dependent on the context.

For instance, solutions at the individual level consist of restructuring the payoff functions for users or by devising selective incentives based on the motivations of the users. This can be accomplished by either reducing the perceived costs or increasing the perceived benefits of contributing. The second type of solution at the individual level focuses on increasing perceived efficacy of individual contributions. Individuals are more willing to participate if they believe their contributions will be valuable to others.

In order to increase individual incentives to exchange knowledge, organizations can explore either of these two options. The first solution would be to selectively reward individual contributions by means of some kind of participation-contingent compensation. The reward does not have to be monetary (although in corporate KMSs it can). One can design non-monetary rewards, such as social recognition, can be extremely powerful incentives as long as they are public, infrequent, credible, and culturally meaningful[77]. However, the design of such incentives also carries a cost – from the monitoring of participation for administering the rewards, the rewards themselves and the probability that it might yield quantity at the expense of quality.

Most discussion on incentives ignore quality of content contributed. The issue of quality vs. quantity often comes up in discussions on throughput in academic knowledge networks that rely on highly qualified content like GSSD. As Carbera[78] points out- a highly subsidized system might encourage employees to artificially inflate their number of contributions at the expense of disregarding their quality. So, while the subsidies may well succeed in increasing cooperation, they can also jeopardize the value of the ideas being shared, ultimately undermining the sustainability of the network.

The solution is to devise incentives that would also incorporate the quality of contribution. This has proved very difficult to do in practice – the use of collaborative filtering, rating of documents by users, rating users etc. has run into problems because second level incentives will have to be devised to make participants contribute to the rating process itself.

---

[77] Lawler, Edward E. III (2000) Rewarding excellence. San Francisco, CA: Jossey-Bass.
[78] Ángel Cabrera,Elizabeth F. Cabrera: Knowledge-Sharing Dilemmas, Organization Studies, 2002, 23 (5): pp 687-710

Perhaps a better strategy would be to increase the perceived value of the collective gain. Although probably more difficult to achieve, this approach has the advantage that employees will be more likely use their best judgment concerning their contributions in order to maximize collective gains. In corporate KMSs, this amounts to increasing the value of the collective gain by combining the knowledge exchange program with a gain-sharing or profit sharing plan – where every employee receives a bonus based on the success of the knowledge-sharing program. The reward thus depends on the combined efforts of the individual and the other people with whom he or she exchanges knowledge. Given that the employee's contributions should benefit the work of other group members, these contributions should also increase the potential value of the gain-sharing bonus the employee will receive.

There is a clear difference between this solution and that of individual selective incentives. In this case an employee is not rewarded directly for contributing to the shared database. The motivation to contribute comes from the fact that the employee will receive a reward if the knowledge-sharing program is successful.

Solutions geared towards gathering critical mass and getting to the tipping point would have to realize that in large communities, participants may perceive that their contributions do not make a big difference. Also, homogenous larger groups may lead to higher redundancy rates in the repositories, thus harming both the probability of participants making valuable contributions and the probability of users finding what they need. This again brings us back to the issue of quality vs. quantity, although both of these problems, increased redundancy and search difficulty, could be solved through the adoption of advanced technology(better search/indexing techniques, navigational aids etc.)

The use of institutions to correct collective action bring with them the cost of establishing organizational structures, enforcement mechanisms plus the added welfare loss associated with moving away from market based solutions. Also there is the question of who will guard the guardians. Experience has shown that electronic knowledge networks that impose non-contractual sanctions, informally or through a governance structure, is costly and the same free-rider problems that impede cooperation in pursuit of the original collective goal will also impede cooperation in enforcing norms, creating a second-order collective action problem.

Our discussion so far has focused on the theory behind knowledge networks and the problems these networks in achieving sustainability, analysing the motivations of people who contribute to knowledge networks, and suggesting what incentives could be used at the individual, network and the institutional level to spur people to contribute. The next phase of this thesis deals with two

important knowledge networks in the real world- the MIT developed GSSD and the Development Gateway's Topic Pages. We shall attempt to use some of the insights from our above discussions to asses the health of these networks, identify barriers towards sustainability and suggest corrective measures.

4. CHAPTER IV: Knowledge Networks in practice (I) : The case of the Development Gateway

4.1. Introduction:

The DG Foundation is head quartered in Washington DC and it began operations in July 2001. Initially conceived and developed in the World Bank, the Development Gateway Foundation aims to improve people's lives in developing countries by building partnerships and information systems that provide access to knowledge for development[79].

The Foundation is currently governed by a Board of Directors composed of directors from 11 countries (Australia; China; Germany; India; Italy; Japan; the Republic of Korea; Luxembourg; Mali, sponsored by the Netherlands; Pakistan; and Rwanda), two businesses (MphasiS and Transnational Computer Technology), the United Nations, and the World Bank Group. The Board is chaired by Dr. Mamphela Ramphele, Managing Director of the World Bank.

Through its various programs the Gateway aims to serve the needs of civil society, the public and private sector and the official donor community by exploiting innovative and affordable information and communication technologies (ICT). The DG foundation runs the Development Gateway Portal that provides a space for communities to share experiences on development efforts.

4.1.1. Objectives and Mission:

The Development Gateway was brought about to respond to the growing recognition that sustainable development requires a long range approach based on access to relevant knowledge, effective and transparent participation of every sector, and country ownership and direction of development agenda. The Development Gateway tries to achieve this by bringing together diverse resources and partnerships to build information tools and services that are useful to development, thus becoming an integral part of the development process itself.

The primary goals of the Development Gateway are to
- Increase knowledge sharing;
- Enable aid effectiveness;
- Improve public sector transparency; and
- Build local capacity to empower communities.

---

[79] Development Gateway Foundation Business Plan, Fiscal Years 2004 to 2006. Available at: http://www.developmentgateway.org/download/217729/Development_Gateway_Business_Plan_15Sept03.doc

Provide information, training and applications

Share experiences and innovations

Have them packaged, organized, and distributed

Development Gateway enables Partners and Users to:

Build local capacity and improve effectiveness

Find additional information, collaborators, opportunities

Apply knowledge and provide services and training
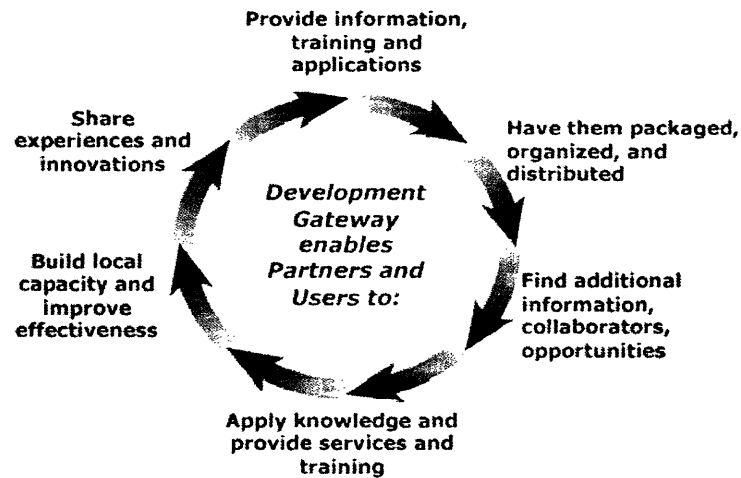
Figure 5: The Development Gateway Dynamic [80]

4.1.2. Service Offerings:

The main services the DG portal offers are-

Topic Pages: To help improve collaboration among development practitioners and facilitate the exchange of information, the Gateway's Topic Pages build interactive communities around key development topics. This knowledge provided by the community is used to increase the impact of development assistance by providing practitioners the right information and contacts to use it effectively.

Through the Development Gateway portal's "Topic Pages," organizations and individuals from over 70 countries share their experiences and knowledge to address key issues and opportunities in development. Expert guides and global topic advisors work with their communities to highlight relevant and useful resources available on the Internet. The Topic Pages have a wide range of knowledge resources on a variety of development topics from Afghanistan Reconstruction to Youth for Development.

AiDA (Accessible Information on Development Activities): The AiDA database is the largest online source of information on development projects in the world. AiDA allows access to over 450,000 historical and current records on planned, current, and completed activities and programs from more than 200 development agencies. Collecting information from many sources,

---

[80] (Source: Development Gateway Business Plan FY04-06

the AiDA database has the potential to become a powerful tool in making aid more effective, improving donor coordination and fostering transparency by presenting its users with a unified interface to view programs and projects by country, region, sector and thematic area.

AiDA partners include governments, bilateral and multilateral donors, development banks, financial institutions, the private sector, foundations, and NGOs. These partners maintain their activity information on their own infrastructure and provide access to these databases by sharing core metadata with AiDA. AiDA's current partners are Bellanet, OECD-DAC, and the World Bank Group. AiDA's information partners use standards for information sharing, thus facilitating the expansion of the database and reducing duplication of efforts.

dgMarket: dgMarket is a global online marketplace providing e-tendering and e-procurement solutions on donor and government-funded development activities. dgMarket aims to increase government transparency and efficiency, and promote business opportunities for local suppliers of goods and services. The global dgMarket tendering service provides access to more than 30,000 current public procurement notices on any given day. Its turnkey e-tendering solutions are also provided to partners and governments for local tendering services in developing nations. dgMarket fosters government transparency and efficiency and contributes to local economic development by enabling small- and medium-sized enterprises to participate in procurement opportunities worldwide. The Development Gateway and the Government of Italy have also established an E-Government Grants Program to rapidly develop and deploy promising e-government applications and activities in developing countries.

Country Gateways: The Development Gateway Foundation's network of 58 Country Gateways are locally-owned partnership-based initiatives that pursue the Foundation's mission on a local level. Country Gateways assist in creating and sharing local and global development knowledge, solutions, and opportunities, encourage business opportunities, and increase the country's exposure.
In countries from Argentina to the West Bank & Gaza, Country Gateways build local capacity, improve access to knowledge, and foster innovation for local economic and social development.
Built on transparent and broad based partnerships, Country Gateways usually consist of a country-level portal and a range of online and offline initiatives that effectively apply information and communication technologies (ICT) to reduce poverty and promote sustainable development.

The programs and services of the DG Portal support the goals of the foundation - dgMarket and the E-Government Program aim to improve public sector transparency; the AiDA database and the Development Gateway portal enable aid effectiveness; the topic Pages of the Development

Gateway portal increase knowledge sharing; and the network of Country Gateways, the Development Gateway portal and dgMarket build local capacity to empower communities.

4.1.3. Organizational Structure

The Development Gateway portal is incubated by the World Bank, and managed by the Board of Directors of the Development Gateway Foundation. An Executive Committee of four to six members, appointed by the Development Gateway Foundation Board and under an elected chair, meets frequently to address governance issues of the Development Gateway portal, to approve grants and to supervise the Development Gateway Foundation's Secretariat.

In addition, the Development Gateway Portal has an Editorial Advisory Committee that provides guidance on matters primarily related to content. The committee examines and proposes ways to enhance participation by communities and experts to promote the highest possible quality of content.

The Development Gateway is led by a Director who is responsible for managing development and delivery of the portal. A core team of 25 full-time staff assists him. In FY01, the World Bank contributed a total of $7.0 million as an investment in the Development Gateway Foundation to help implement the start-up costs of the development Gateway portal.

4.2. Content and Traffic at the Development Gateway:

Traffic analysis at the DG Portal was carried by the author after examining the server log file data and using DeepMetrix LiveStats analysis software for the period Sep 2002 –July 2003. Some of the reports reproduced below are available at the DG Business Plan- summary figures.
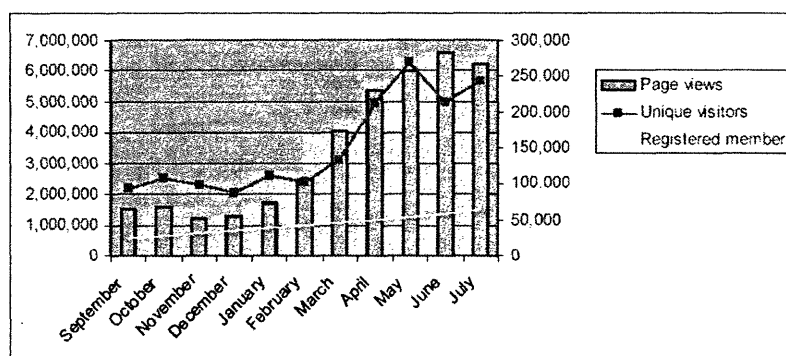
Figure 6: Traffic at the Development Gateway portal[81]

---

[81] Source: Development Gateway Business Plan FY03-04 p 3

There has been a 200% increase in traffic (in terms of visits) to the DG Portal over the last year, with over 500,000 monthly visits and with over half of those being unique visitors. This is encouraging as it reflects the Gateway's ability to attract new visitors to the portal. However the high number of unique visitors may partially hide the fact that the number of 'repeat' visitors is on the decline- indicating that existing members are logging on fewer times in a month.

However, there are complicating factors that prevent us from making this inference. One of these is that members of the portal can opt to subscribe to newsletters that summarize the new resources added on to the portal, and thus their participation will not be recorded on the web server log file data. Also filtering for search engines and 'bots' that regularly hit the web server is not yet perfect, and this might result in a downward correction in the 'visits' figures.

Visits divided by service area are-

Service Area	Visitor Sessions
dgMarket	284,090
Ideas and knowledge[1]	123,134
AiDA	17,791
Country Gateways pages[2]	8,781

Figure 7:Traffic by Service Area[82]

Content Throughput Statistics:

❏ Approximately 30,000 Content items – including links submitted to the topic pages, discussion and comments, job postings and events.

❏ Over 450,000 funded activities from 43 organizations in AiDA.

❏ Over 30,000 active tenders from 153 countries in dgMarket.

❏ Information and resources contributed from individuals and organizations in 70 countries.

User Profile Statistics:

User profile statistics are gathered at the time of registration. The table for the year '02-'03 below indicates that almost a third of members are affiliated to the Private sector. This is probably because a large portion of the dgMarket members are small and medium sized private companies. The DG Portal has a single registration form for all its services, and it is not possible to determine affiliations by service area. NGOs, Government/Public Enterprises and Policy

---

[82] Source: Development Gateway Business Plan FY04-06 p4

Research organizations – institutions most likely to be likely to be direct policy influencers, together amount to almost 40% of membership.

Specified affiliation	Members	% of members
Private sector	4,861	27.7%
NGO / Civil Society	3,697	21.1%
Other (specified by member)	3,087	17.6%
Educational institution	2,784	15.9%
Government / Public enterprise	1,936	11.0%
Policy / Research institution	932	5.3%
News / Media outlet	245	1.4%
Total	17,542	100%

Figure 8: Professional Affiliation of Development Gateway users[83]

Membership across developed and developing countries, as seen by the table below, are about equal. However, what seems to be intriguing is that *the percentage of membership seems to be inversely proportional to the level of development.* This is most likely to be due to a combination of factors- total number of people living in each of these countries, the level of penetration in such countries, dgMarket or Country Gateway networks in their region, the prevalence of other sources of information(country owned knowledge networks etc.) in higher income developing countries, language issues etc.

		Members	% of total
Developed countries	High income	33,819	53.2%
Developing countries	Upper-middle income	7,507	11.8%
	Lower-middle income	10,702	16.8%
	Low income	11,545	18.2%
Total		63,573	100%

Country classifications from the World Bank's World Development Indicators

Figure 9: Source of Development Gateway traffic[84]

The graphic below gives one a better picture of user traffic by continent.

[83] Source: Development Gateway Business Plan FY04-06 p4
[84] Source: Development Gateway Business Plan FY04-06 p6
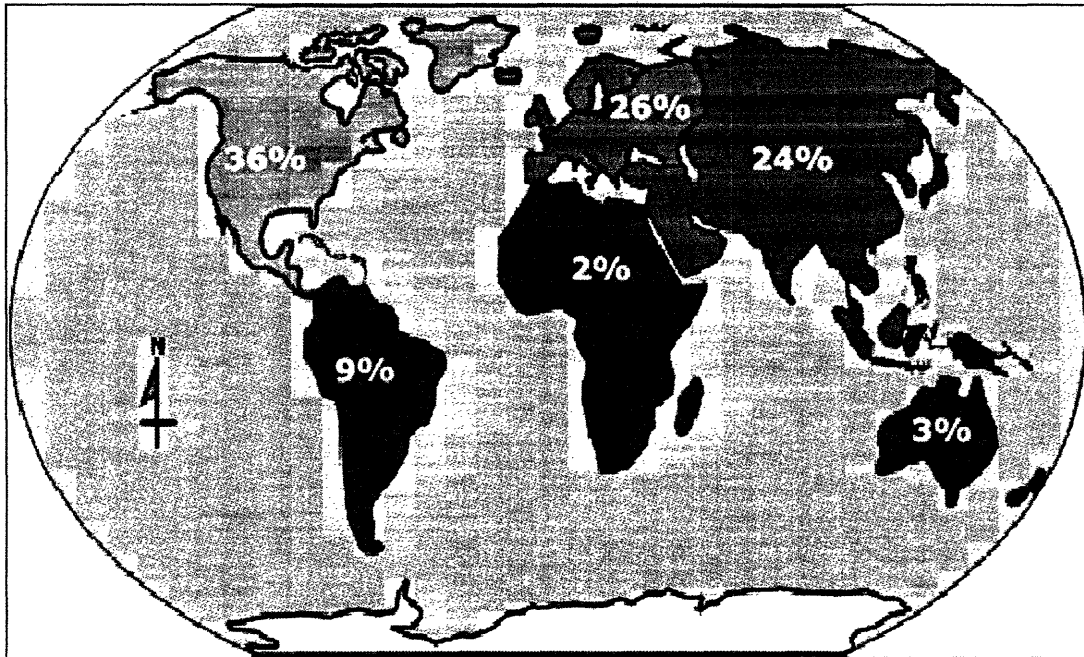
Figure 10: Traffic by continent[85]

4.3. Knowledge Networks in the Development Gateway:

The mandate of the Development Gateway at the Global Level is to build information tools and services for the provision of the public good. The portal services provide global access to information, experiences, and resources on topics, projects, development activities, and procurement opportunities. The Gateway takes the view that sharing knowledge is one of the keys to building local ownership. When people feel fully connected to the relevant knowledge in their domain, they become more effective proponents of their own agenda for development. This is also reflected in their approach to local capacity building.

At the local level, the mandate is to build individual, institutional, technical, and economic capacity through the Country Gateway, AiDA and the dgMarket portals by providing knowledge that would otherwise unavailable to many individuals doing development work in the field.

The portal supports the Development Gateway Foundation's knowledge-sharing efforts through services such as an online directory for information on development projects (AiDA), an electronic

---

[85] Source: Development Gateway Business Plan FY04-06 p7

procurement market (dgMarket), information on major development topics (Topic Pages), and links to a growing network of country-level initiatives (Country Gateways).

The Development Gateway's unique value to the development community comes from the combination of three distinct approaches:

Decentralized Community Model

The Gateway has adopted a decentralized community model that provides for local and distributed site ownership, content management and administration. This facilitates user communities and Country Gateways to develop their own sites, while sharing the Development Gateway platform. The decentralized model provides communities and working groups with tools that enable them to locally brand and determine the content that appears on these pages, and can modify the appearance of the pages to give their sub-site a unique look and feel. The team is working closely with several communities on pilot projects to further identify the necessary tools.

Content Management Partnerships

Content drives the Gateway. It provides the "value added" to audiences in developing countries and to other stakeholders. To provide high-quality content, the Gateway partners with organizations and companies willing to take ownership of, or co-brand, a development topic. The Gateway provides the tools and platform to its partners to build topics and expand their networks, share knowledge, and engage the development community in working toward common objectives.

Open Technology and Information Standards

The Development Gateway has pursued an approach to technology and information standards, which enables greater accessibility, scalability and collaboration. An open-source technology platform has been adopted. The Development Gateway also builds common approaches and standards within the development community on the design, implementation, and management of information-sharing activities on development.

4.3.1. Partnership Types:

- **Content Partners**: Managing a content site on the Gateway platform. This would allow an organization to share its expertise in a key sector, region or country while expanding its reach to a broader audience.

- **Country Gateway**: Establishing Country Gateway organizations to become the domestic umbrella organization(s) in developing countries. These organizations would develop community sites on the Development Gateway for private sector, civil society, or government partners to strengthen local community development efforts while providing valuable information and experience through the common platform to a worldwide audience.

- **Hosting Partnerships**: Co-hosting or co-branding commercially viable development solutions with the Development Gateway that bring valuable products to both business and governments while providing revenue stream to the project.

- **Development Partnerships**: Participating in the Gateway's integrated development project database to build value-added information across the community on issues such as aid matching and foreign direct investment.

- **Other Partnerships** like seconding skilled staff to the Gateway to contribute valuable skills while gaining experience in a unique large-scale development initiative.

4.3.2. Information and Technology Architecture:

The Development Gateway is a database-backed web environment that enables decentralized content management. The technology architecture is based on the following design principles:

- Use of open-source and standards-based technologies;

- Distributed content, coordination and ownership;

- Sharing of the Development Gateway technology.

An important component of the Development Gateway architecture strategy is the use of open-source software. The Development Gateway was implemented through the use of the ArsDigita Community System (ACS), an open-source code software system for interactive Web applications. With the exception of the Oracle database, all components of the Development Gateway platform are open-source code, which enables it to transfer the technology and encourage partners to become co-developers of the Development Gateway platform. The portal currently runs on an AOL server that is also available on the public domain. In addition, extensive development and use of XML standards allows participation regardless of the platform of the partners.

Although the Development Gateway Portal can be viewed as a knowledge network in itself, offering a variety of programs and services for knowledge sharing and dissemination by providing an interactive space for communities to share experiences on development efforts, our focus here is on the Topic Pages program of the Development Gateway. The presence of different kinds of knowledge networks with in the portal indicate an adaptation to the content and the structure of the network to the reality of its context. In this respect, there are some distinctive differences between international (topic pages) and local (country gateway) networks. The strengths of international networks are in promoting ideas, providing technical support, coordinating policies and disseminating products – which the topic pages are geared to provide for. Local country gateway networks are more conducive to capacity development and project collaboration.

4.3.3. The Topic Pages as a Knowledge Network:

The topic pages by offering users virtual interactive communities centered on development issues encourages knowledge sharing and helps improve collaboration among development practitioners. The objective of the topic pages is to offer easy access to *relevant, high quality,* and *customized information* from *diverse sources.*

4.3.4. Content Structure and Organization within the Topic Pages

To ensure **relevancy** the topic pages the Development Gateway concentrates its knowledge sharing efforts into to 35 subject areas on a wide range of topics related to sustainable development. The topics are reproduced below-

Afghanistan Reconstruction	ICT for Development
Aid Effectiveness	Indigenous Knowledge
Argentina: Challenges and Opportunities	Indigenous Peoples
Business Environment	Indigenous Rights
Capacity Development for MDGs	Innovations for Development
Civic Engagement	Iraq: Relief and Recovery
Culture and Development	Judicial and Legal Reform
E-Commerce for Arts & Crafts	Knowledge Economy
E-Government	Microfinance
E-Learning	Non-Governmental Organizations
Environmental Law	Population and Reproductive Health
Food Security	Poverty
Foreign Direct Investment	Privatization
Gender and Development	Trade and Development
Glocalization	Urban Development
Governance	Water Resources Management

85

Figure 11: List of all topics in the Development Gateway

Within each topic, content is further categorized as belonging to either of several resource types-

- News
- Data and Statistics
- Documents and Reports
- Events and Discussion Forums
- How to / Tools
- Organizations, Networks, People
- Programs and Projects
- Publications and Multimedia

The most frequently submitted resources are links to other websites and documents that are relevant to the topic. Content is also submitted through comments on resources, bulletin boards and discussion groups. Each topic page also offers other relevant content to its users like a calendar of events related to the topic, jobs, related publications etc. Topic pages allow for subscribers to receive weekly/monthly newsletters and the notification of alerts when topic relevant to them is added or commented upon.

Content Management within the Topic Pages:

To ensure quality of the resources submitted, only registered members can submit resources to the topic pages. Unregistered users can however view most topic resources, but are not eligible to receive subscriber alerts and newsletters.

A subscriber is a registered user of the DG portal. A subscriber thus subscribes to all services in the portal. A subscribing member (sometimes referred to as just member) is a subscriber who expresses interest in a Topic Page at the time of registration. Thus each subscriber can be a member of one or more topics within the topic pages.

Members	Number of members	% of total
Browsing members	32,676	48.4%
Subscribing members	34,814	51.6%
Total – all members	67,490	100.0%
Newsletter subscribers	8,849	13.1%
Newsletter + topic subscribers	15,898	23.6%

Figure 12: Usage statistics by membership type across all Topics

The above table indicates that close to half the people who view resources on the topic pages are unregistered/unsubscribed users.

Quality Control and Moderation Processes within the Topic Pages:

The overriding goal of the Gateway is to ensure an *open, pluralistic and quality controlled content management system*. The content management system of the site is based on the editorial judgment of selected individuals and organizations called **guides,** chosen based on their knowledge of the subject, objectivity, and openness. Guides are responsive to direction from **advisors,** also experts in the subject area, who themselves are drawn from a wide range of institutions and stakeholders.

As mentioned before, only subscribers can contribute resources to a topic. Content submitted to the Gateway by users is subject to prior review by editors or guides who determine its suitability for inclusion. Gateway guides or editors are under no obligation to link to any site, or to accept the contribution of any individual or organization and the sole criteria for inclusion is the usefulness of the resource for the intended audience. Links and resources are periodically checked for suitability.

Eight of the thirty five topic guides are full time employees of the Development Gateway, while the others are from the World Bank, NGOs, academic institutions or are private consultants.

The topic pages use a 'deferred publishing'[86] approach, so that content suggested by users is reviewed prior to acceptance on the site. Selection of resources, issues, discussion topics, news, and technical data for the site is, to the extent possible, based on fact, careful analysis, and the needs of the Gateway audience. This challenge of coupling Internet technologies with sound

---

[86] From the Editorial Policy of the Development Gateway: Available at
http://www.developmentgateway.org/EditorialPolicy

human judgment -- in a manner that is both open and professionally sound -- is at the core of the Gateway's editorial policy. However the sound human judgment provided by the gateway through its guides, advisors and editors comes with a significant labor cost.

The Editorial Committee provides guidance to the Development Gateway Foundation's Board of Directors for assuring quality on the Gateway portal. Its role is to advise the Board on matters such as systems of content management, editorial policies, publication standards, and content partnership structuring.

4.4. Problem Definition

In April 2003, a report prepared by an independent consultant on the current status and accomplishments of the Development Gateway[87] identified the core issue with the Topic Pages as "... *The cost of provision and quality assurance of Content information grows in relation to the amount of information provided, making it difficult to balance the need to provide more and better free information with the available funds.*"... *"The Development Gateway needs to change the nature of this relationship.."*

Currently, the Development Gateway bears almost all of the cost of content provision and quality assurance (QA). The topic guides are responsible for most of the content that flows through the network, with subscribers contributing very little in terms of new content or comments on existing content. The Gateway's approach so far has been to recruit new topic guides or increase the breadth of current topic guides coverage of each topic guide. This model clearly does not scale.

In only its third year of operation, the Gateway has recognized this fact, and is looking for new models of participation, where content provision is delinked from the costs of running the network. As we have seen before, the answer to this problem involves creating sustainable participation models where the community is incentivised to contribute to the public good. The problem thus resolves itself to one of analyzing the motivations of members, using these to create the right incentives for participation, and embedding these incentives into the technology, processes and institutional structure of the knowledge network.

---

[87] Prepared by: Louise Walker Consulting, Louise Walker, Heidi Hay: Available on the Development Gateway website at
http://www.developmentgateway.org/download/220436/DG_Accomplishments_Final.pdf

Another sub-problem we shall attempt to solve is the resource allocation problem of how best to allocate 'guides' to 'topics' using the Critical Mass theory.

4.5. Content Provision Patterns within the Topic Pages:

Content Throughput, User Registration and Web Traffic data for the following 11 topics were collected: **"Afghanistan Reconstruction", "Aid Effectiveness", "Civic Engagement ", "E-Government", "Gender", "Governance", " ICT for Development", "Iraq: Relief and Recovery", "Non-Governmental Organizations", "Poverty", "Trade and Development".** This is a representative sample that takes into account all levels of user activity (high medium low), type of topics (issue based/event based) and newly introduced vs. old topics.

The data was culled from the Development Gateway Portal's database on user registration and content management system, and web server log files.

The data collected from User Registration and the Content Management System included the following items:

- ✦ **New members:** Number of registered new members who joined that month
- ✦ **Total resources viewed:** Includes all resources (links, news, events, specials, newsletters)
- ✦ **Accesses to news:** Number of visits to the 'News' Section
- ✦ **Accesses to events:** Number of visits to the 'Events' Section
- ✦ **Resources viewed by registered members**
- ✦ **Resources viewed by non members** (also referred to as 'browsing' members)
- ✦ **Resources contributed:** Refers to all contributions including links, 'specials', discussions etc. from both guides and members
- ✦ **Resources contributed by members**
- ✦ **Number of members who contributed** –Refers to only registered members, as anonymous contributions are not possible.
- ✦ **Resources published:** Equal to the total resources submitted minus the number of resources rejected by the guides
- ✦ **Searches within a topic page:** Web based searches on resources within the topic
- ✦ **Broadcast send:** E-mail broadcast sent by members

⤺ **Resources mailed to others:** Number of resources mailed by members to others (Each resource is provided with a 'mail to:' link, which lets users mail a resource to anyone within or without the Gateway)

4.5.1.   Assessing Contribution Level by Topic:

A plot of total contributions/ total membership across all topics indicates that some topics are clearly underprovided with respect to others. At this stage it has to be noted that no mention is made of the quality of resources provided. This is a conscious decision due to the fact that 1.it is impossible to objectively define quality of contributions (as this brings forward issues like quality to whom, and how to measure quality – on a stand alone basis based on the content, the ultimate value it provides to the user, or the use the knowledge it is put to by the user, etc.) , and 2. the underlying assumption that any resource deemed to be suitable to publish automatically becomes one of 'acceptable' quality.

To determine the relationship between the number of resources contributed to the number of members subscribed to a topic, a correlation analysis was performed which yielded a moderately[88] high positive correlation coefficient of 0.47.
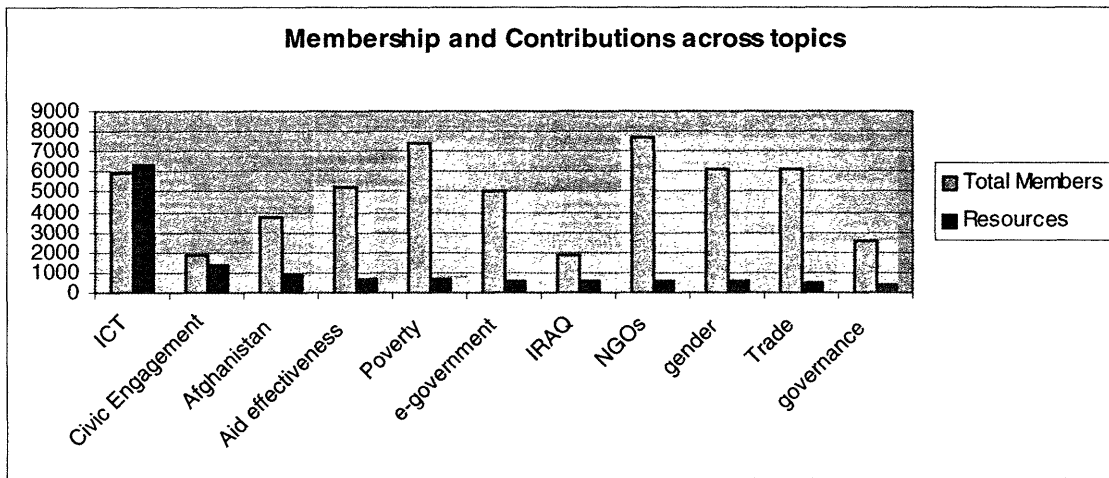
Figure 13: Membership and contribution by topic area

A further analysis looked at the total contributions/ total membership across topics and it was found that for the top 3 topics in terms of Resources contributed, the contributions/ member are also the highest.

---

[88] Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). New Jersey: Lawrence Erlbaum.
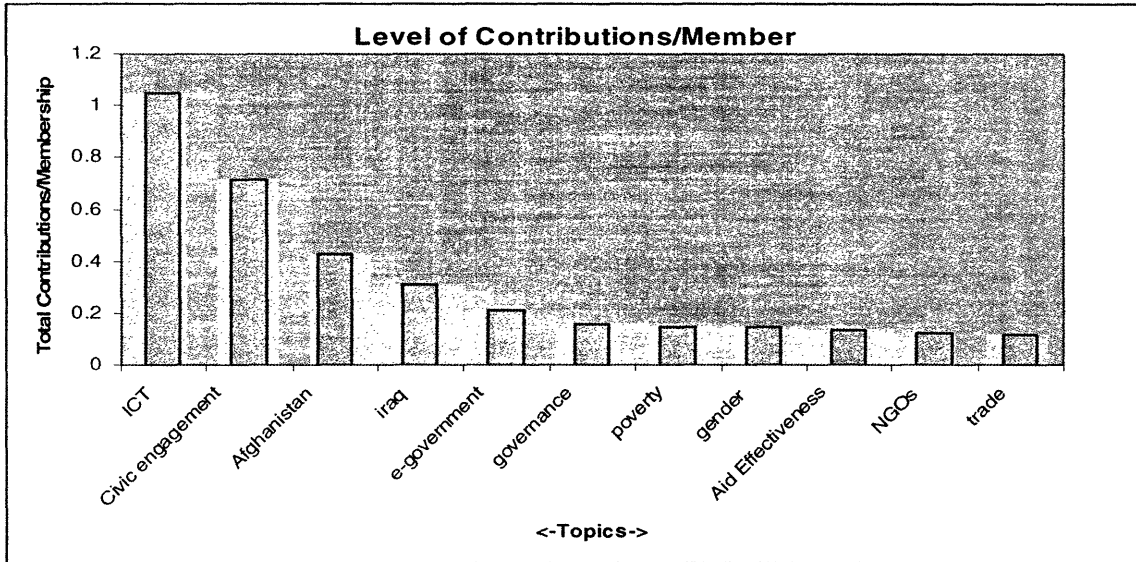
Figure 14: Number of resources contributed/member across topics

4.5.2.   Extent of Community Participation:

The level of contributions in the above cases takes into account contributions by both topic guides and by other members/subscribers of the topics. Plotting member's contributions leaving out the resources contributed by topic guides) gives us a better picture of the *community's* contribution to the topic pages (although topic guides are an integral part of the development gateway community, the motivations and the incentives driving them are different from the others).
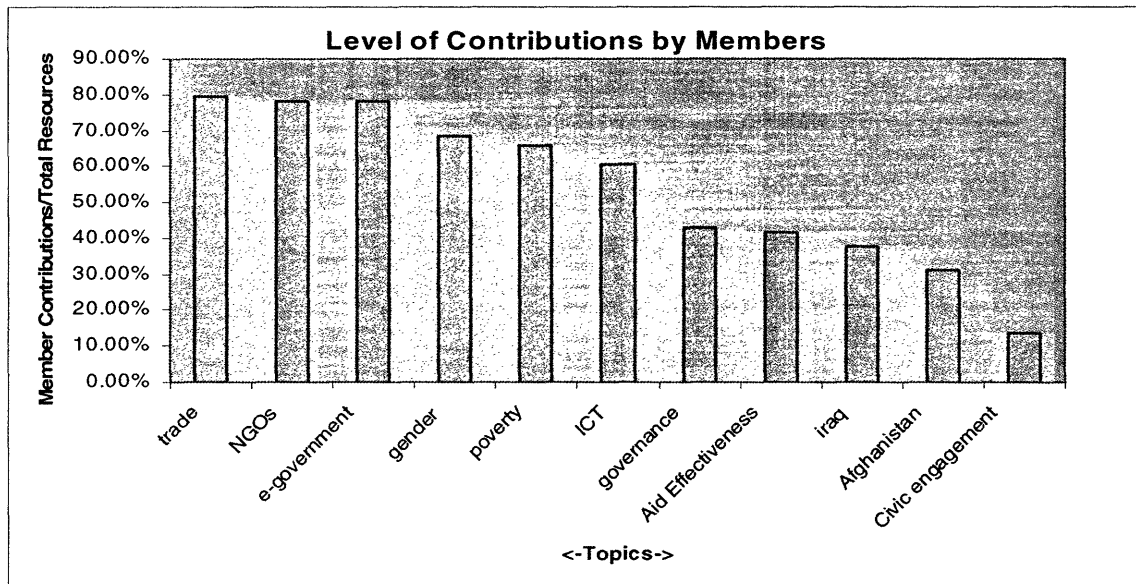
Figure 15: Extent of community participation in resource provision

The above picture plots the community's (non guide contributions) as a percentage of total resources contributed across topics. We see that 'Trade', 'NGOs' and 'e-government' rank the highest as more than 75% of the resources contributed come from the community.

From the above graph and from calculating the correlation coefficient between community contributions vs. total resources contributed (corr. coefficient -.01) there seems to be no apparent aggregate relation between how much the community contributes and the level of resources provided within a topic. Indeed, *Trade, NGOs and e-government that account for the last three topics in terms of total resources provided rank among the highest three topics where community participation(as a fraction of total resources provided) is the highest.*

One notable exception to this phenomenon seems to be the ICT topic- the largest (in terms of resources provided) and the longest standing topic has a healthy 65% of its resources coming from community contributions.

4.5.3. · Impact of contribution heterogeneity on Content Provision:

An analysis of how many members actually contribute as a fraction of the overall membership is an indicator of the *heterogeneity of content provision* and provides an insight into the extent of involvement of community participants. Although data is not available on how many people contributed how many resources within each topic (it is expected that this distribution would follow a Zipf's or an equivalent power law), aggregate statistics on how many members contributed *at least* one resource each month is available. This statistic was averaged over the period of study (Nov 2002-2003). The results are reproduced
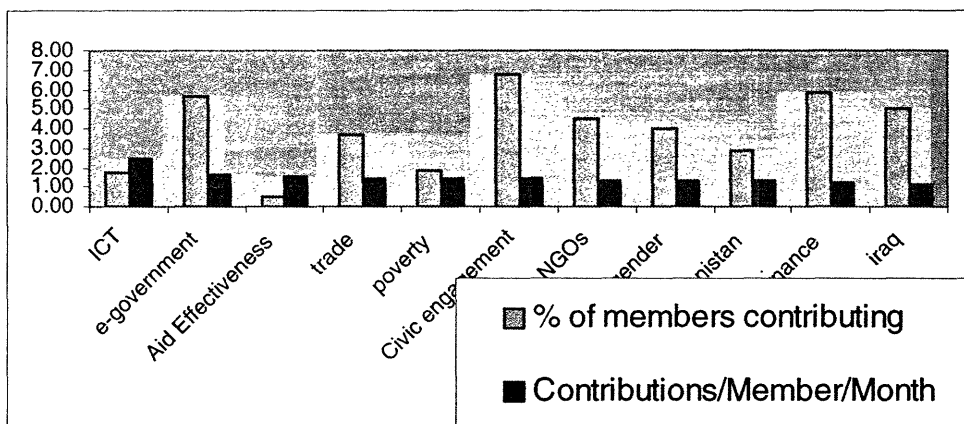
Figure 16: Fraction of members who contribute

The correlation coefficient between number of resources contributed by members and the percentage of members who contributed at least one resource is 0.7. However, correlation does not imply causation and one cannot definitively infer that a heterogeneous mix of contributions **leads** to more content provided without looking into the distribution or the mix of contributions *within* each topic. Other factors that prevent us from making this inference include the differences in topics themselves – the age and maturity[89] of the topic, user demographics, the kind of topic etc.

Contributions/contributing member each month seems to be on an average of 2 resources every month across all topics and is very strongly related to over all resources within a topic. However, this statistic is not entirely reliable and is to be empirically tested as it is expected that a few contributors would provide the bulk of the resources.

4.5.4.   Impact of Moderation and Quality Control:

The impact of quality control and the moderation process on the level of resource provision was then looked at. Specifically, we wanted to test if the 'reject' percentage had any relation to the total level of contributions. The results of this analysis are reproduced below-
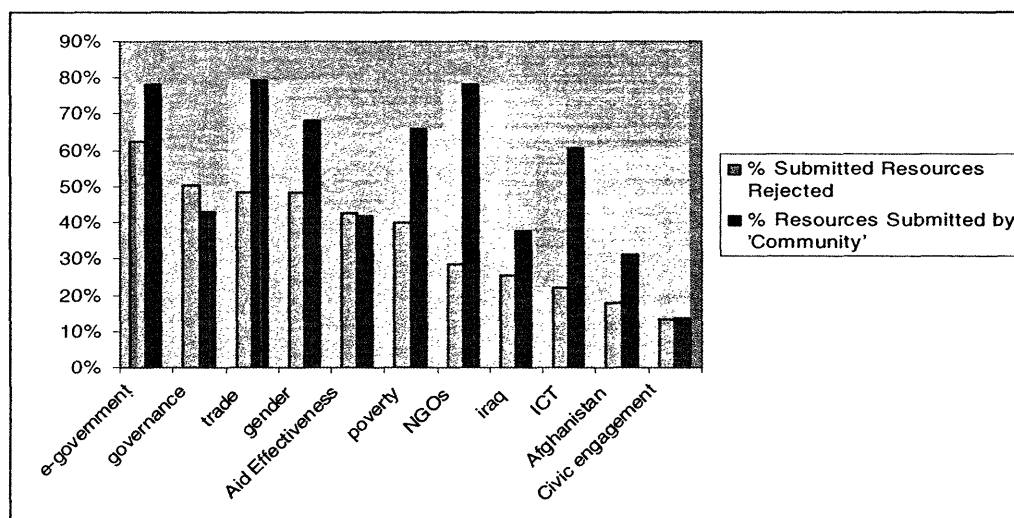
Figure 17: % age of submissions rejected

---

[89] Mature topics as defined here are those where the rate of growth of membership, contributions and percentage rejects is more or less stable. This is mostly a function of the age of the topic- topics like ICT, Aid effectiveness which were among the first few topics to be incorporated.

The reject percentage as calculated here is the difference between the resources published and the resources contributed in a month. Although the entire difference cannot be attributed to rejects due to unsuitability of content (as resources submitted at the end of the month would be published next month) this statistic still provides a good measure to identify the quality of contributions to a topic. e-government, governance, trade and gender have more than 1 in 2 resources rejected- an extremely high statistic.

The above chart compares the % of resources which are rejected with the total number of resources submitted to that topic. There seems to be an inverse link (corr. coeff. -0.52), meaning that the reject percentage seems to be lower for topics with more resources – indicating topic maturity and learning effects within the topic over time. Also reject percentage increases with community participation (put differently, reject percentages are lower when guide participation is high- an obvious result). Also in more mature topics- ICT/ Aid effectiveness, reject percentage seem to be plateauing.

The following table shows the computed correlation coefficients between the most popular topics(in terms of resources viewed) and the number of resources published and the number of members within that topic. Clearly, the more popular topics are the ones with the most published resources. However, the link between membership and total resources viewed is not that strong – indicating a higher percentage of 'browsing' members (48% from fig. 11-7)

	Total Members	Total resources viewed	Resources published
Total Members	1		
Total resources viewed	0.406844	1	
Resources published	0.43995	0.815219	1

An analysis of contributions over time for each topic was carried out. Data was taken for a 23 month period from Jan 2002 till November 2003. While some of the topics were introduced in Jan 2002 (ICT, Aid effectiveness etc.), some of them were introduced only in 2003(Civic engagement, Gender, Iraq etc.)

As can be seen from the figures below, in two of the oldest topics- the growth in membership contributions seem to be plateau and/or decline with time
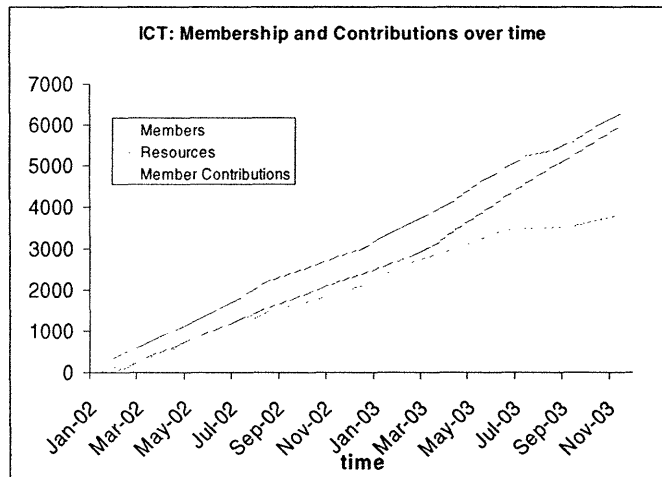
Figure 18: ICT membership and contribution over time
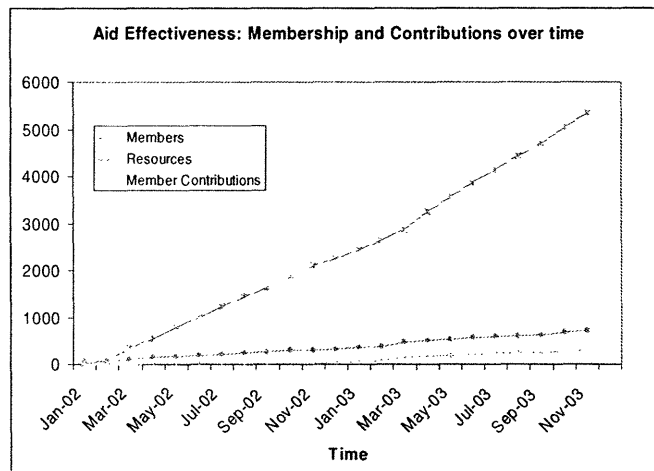
Figure 19: Aid effectiveness, membership and contribution over time

4.5.5.   Community Characteristics:

To test if the characteristics of the community had any relation to the extent of resources contributed, the two available demographic variables- geographical location of the user, and the user's professional affiliation were examined. The two figures below show us the nature of the community participating in the ICT topic.
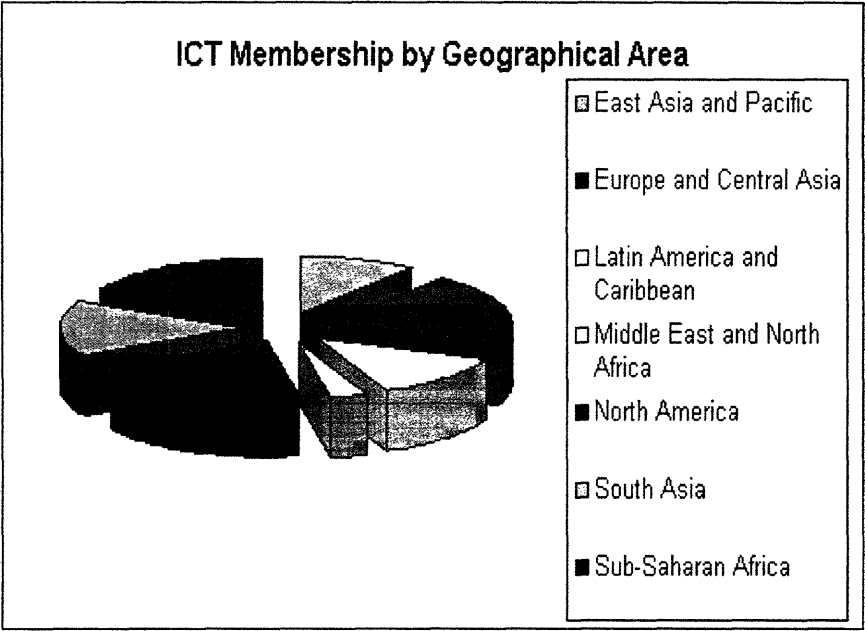
Figure 20: Geographical distribution of ICT Members
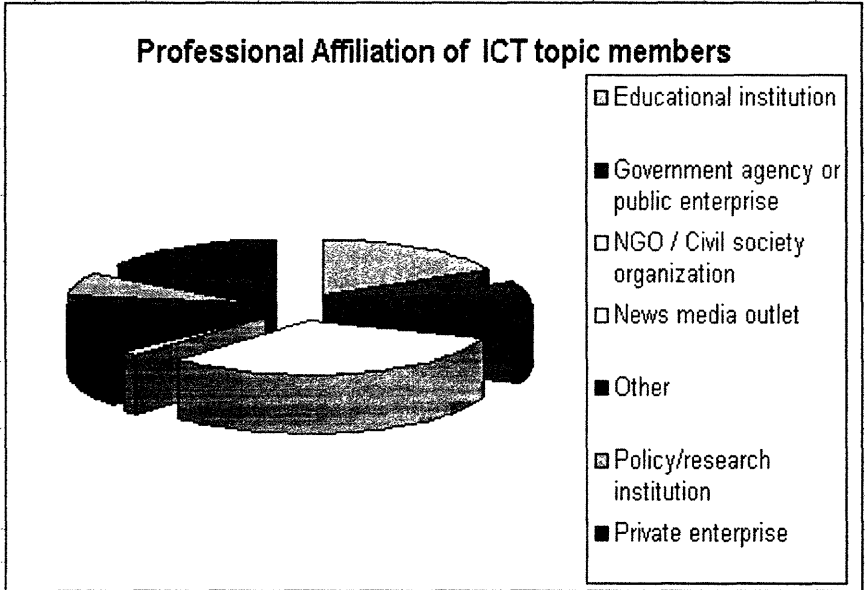
Figure 21: Professional Affiliation of ICT Members

The community's geographical characteristics does not vary from topic to topic to any appreciable extent, although certain topics of high interest to a region have slightly skewed membership patterns (for example- North America and 'Iraq', Latin America and 'e-governance', South Asia and 'Afghanistan Reconstruction etc.). Contributions from the Middle East and North Africa were uniformly low, except for the 'Iraq' topic.

The professional affiliation of members across topics exhibits more variability. Topics that are of obvious interest to the member's organizations such as Government Agencies and 'e-governance' and 'governance', News media outlets and 'Iraq', Private enterprise and 'Trade', 'Afghanistan Reconstruction' and 'Iraq' tend to have more representation from them.

A regression analysis to test if topics with a higher proportion of a particular professional affiliation were linked to higher resource contributions or higher overall membership showed that there might be a slight possibility (regression coefficient 0.45, p value=0.2), that a higher percentage of policy/research institutions in the membership distribution might lead to greater contributions. This result will have to be empirically tested and reproduced with more data from *within* a particular topic.

4.5.6.   Revisiting Critical Mass Theory:

Section 3.4.1 explored the application of critical mass theory in the context of knowledge networks.

Specifically we looked at the incentives to contribute in the early stages of community formation and put forward the theory that it progressively gets easier to entice voluntary contributions once the network grows in size (in terms of members and resources) - the production function for knowledge networks would be accelerating in nature when the members of the network perceive higher levels of activity (either through new members joining, through new resources being posted-by existing members or by guides)

To rephrase, the hypothesis being tested in this section, is whether the level of contribution of resources in any time period is a function of the already existing level of resources or the activity in the preceding period. The theoretical argument is that if we were to control for all other factors (the natural growth/change in the membership in the network due to maturity, advertising and attrition, and assuming the same quality of resources contributed in each time period), the Resources contributed by *members* in time period t $R_c(t)$ would be governed by the multivariate time series equation-

$R_c(t) = a + b_1 R_c(t-1) + b_{11} R_g(t-1) + b_{111} N_m(t-1) + b_2 R_c(t-2) + b_{22} R_g(t-2) + b_{222} N_m(t-2) + b_3 R_c(t-3) + b_{33} R_g(t-3) + \ldots b_{nnnn} R_n(1) + b_{nnnn} R_g(1) + b_{nnnn} N_m(1)$

Where $R_g(t)$ is the number of resources contributed by guides in a time period t, $N_m(t)$ is the number of members who contributed in the time period and all a_i, b_i 's are constants.

The choice of time period t is important. While t should not be too small- it should allow for enough number of members to view the topics that have been contributed in the previous period and act on them, the choice of t should not be large enough to mask the effect of contribution dynamics within a time period. The ideal choice of t would be one which would coincide with the average time lapsed between two successive visits by an average user.

Due to the difficulty in computing this statistic, and the fact that time-series data was only available on a month by month basis, the t value for this analysis is taken as one month, although it should be noted that a t value of a fortnight would probably be a better unit of analysis. It is essential that the results of this section be reproduced with a shorter time period to prove the validity of the critical mass hypothesis.

To explore the effect of each of the contributing variables the *backward elimination* algorithm in subset/stepwise regression was used to find the best subset of dependent variables that would explain the variability of the independent variable R_C (t). Also to account for auto correlation, the results were run with first dropping $R_c(t-1)$ and then $N_m(t-1)$ for each time period. The results of the analysis for the ICT topic are reproduced below-

# Coefficients	R-Squared	Adjusted R-Squared	Models (Constant present in all models)									
			1	2	3	4	5	6	7	8	9	10
2	0.33	0.29	Constant	Rc(t-1)	*	*	*	*	*	*	*	*
3	0.43	0.36	Constant	Rc(t-1)	Rg(t-2)	*	*	*	*	*	*	*
4	0.64	0.57	Constant	Rc(t-1)	Rg(t-2)	Rc(t-2)	*	*	*	*	*	*
5	0.68	0.59	Constant	Rc(t-1)	Rg(t-2)	Rc(t-2)	Rg(t-3)	*	*	*	*	*
6	0.69	0.58	Constant	Rc(t-1)	Rg(t-2)	Rc(t-2)	Rg(t-3)	Nm(t-3)	*	*	*	*
7	0.73	0.59	Constant	Rc(t-1)	Rg(t-2)	Rc(t-2)	Rg(t-3)	Rc(t-3)	Nm(t-3)	*	*	*
8	0.74	0.57	Constant	Rg(t-1)	Rc(t-1)	Rg(t-2)	Rc(t-2)	Rg(t-3)	Rc(t-3)	Nm(t-3)	*	*
9	0.74	0.53	Constant	Rg(t-1)	Rc(t-1)	Nm(t-1)	Rg(t-2)	Rc(t-2)	Rg(t-3)	Rc(t-3)	Nm(t-3)	*
10	0.74	0.48	Constant	Rg(t-1)	Rc(t-1)	Nm(t-1)	Rg(t-2)	Rc(t-2)	Nm(t-2)	Rg(t-3)	Rc(t-3)	Nm(t-3)

Table 4-16: Regression results for ICT

A similar analysis for the Aid effectiveness topic shows the following results-

# Coefficients	R-Squared	Adjusted R-Squared	Models (Constant present in all models)								
			1	2	3	4	5	6	7	8	9
2	0.35	0.32	Constant	Rg(t-2)	*	*	*	*	*	*	*
3	0.52	0.46	Constant	Rc(t-1)	Rg(t-2)	*	*	*	*	*	*
4	0.54	0.45	Constant	Rc(t-1)	Rg(t-2)	Rc(t-2)	*	*	*	*	*
5	0.56	0.44	Constant	Rc(t-1)	Rg(t-2)	Nm(t-2)	Rg(t-3)	*	*	*	*
6	0.60	0.44	Constant	Rc(t-1)	Rg(t-2)	Rc(t-2)	Nm(t-2)	Rg(t-3)	*	*	*

7	0.63	0.44	Constant	Rc(t-1)	Rg(t-2)	Rc(t-2)	Nm(t-2)	Rg(t-3)	Rc(t-3)	*	*
8	0.63	0.39	Constant	Rc(t-1)	Rg(t-2)	Rc(t-2)	Nm(t-2)	Rg(t-3)	Rc(t-3)	Nm(t-3)	*
9	0.63	0.34	Constant	Rg(t-1)	Rc(t-1)	Rg(t-2)	Rc(t-2)	Nm(t-2)	Rg(t-3)	Rc(t-3)	Nm(t-3)

Table 4: Regression results for Aid Effectiveness

Interpreting the Above Tables: The tables are regression outputs that indicate the best subset of variables that can predict the value of the independent variable. The table lists the # of coefficients in the model, the Rsquare and the adjusted Rsquare *for each model*. The models themselves are shown with the independent variables they use and the constant. The best model is the one with the highest Adjusted Rsquare, and these models are highlighted by shaded rows.

A further analysis of the subsets with the highest Rsquare (adjusted) indicates that the two dependent variables with the most significant 'p' values are Rc(t-1) and Rg(t-1) in the ICT topic and Rc(t-1) and Rg(t-2) in the Aid Effectiveness topic. For example, in the ICT topic, the MLR (Multiple Linear Regression) results yields the following-

Predictor (Indep. Var.)	Coefficient	Std. Error	p-value
Constant	217.73	100.52	0.05
Rg(t-1)	0.16	0.22	0.19
Rc(t-1)	0.54	0.26	0.07

Table 5: Predictive capacity of independent variables

Interpreting the above table: The table above indicates the most significant variables in the model derived from the best subset analysis performed before. Within the best subset, the variables **Rg(t-1)** and **Rc(t-1)** are the best predictors of the dependent variables- as indicated by the size of the coefficients and the low p values.

4.6. Interpretation of Results:

Thus far, we have looked at the nature of contributions across topics, the extent of community participation within these topics, the impact of heterogeneity in content provision on the total amount of resources provided, the impact of moderation processes and quality control and tested the evidence for the critical mass theory within a select few topics in the topic pages. It will have to be repeater here, that the above analysis makes no effort to distinguish between the quality of information resources submitted, and all published resources are deemed to be of 'acceptable quality'.

Although the results would have to be reproduced by looking at similar data for all other topics and varying values of the time period, and the results would have to make allowances for differences amongst the topics- the content, their maturity, the proficiency of the guides etc., it is still possible to make a few cautious inferences on the nature of contributions over time in a knowledge network. Here is a summary of the findings so far-

> The relationship between the number of resources contributed to the number of members subscribed to a topic, a correlation analysis was performed which yielded a moderately

high positive correlation coefficient of 0.47, indicating a possible link between level of membership and the number of resources contained within a topic. However the more significant statistic is the high correlation coefficient between the contributions per member and the number of resources within a topic - for the top 3 topics in terms of Resources contributed (ICT, Civic Engagement and Afghanistan), the contributions/ member were also the highest.

Many knowledge networks use membership (levels and growth) as an indicator of the robustness of the knowledge network. The absence of a direct link between membership levels and resources provided shows that higher levels of membership do not always ensure sufficient content and that contributions/member could prove a much better indicator.

➢ There seems to be no apparent aggregate relation between the extent of community contribution (defined as 'non guide' contributions as a fraction total contributions) and the level of resources provided within a topic, indicating that it is not possible to conclude that community driven topics have higher overall resource levels. However this could also indicate that the percentage of resources contributed by the community is higher because the resources contributed by the guides are low, which might be an indicator of the lower priority of the topic/area itself, lack of availability of resources to submit etc. To test the hypothesis that community participation leads to higher resource contributions, one needs to perform the analysis for two 'identical' , mature topics with similar community characteristics. Unfortunately, it is not possible to perform this analysis for any of the topics on the Gateway with the available data.

➢ Heterogeneity of content provision (what percentage of members contribute) is highly correlated with the overall number of resources, probably indicating that diversity in members who contribute could contribute significantly to the overall resource level. The reason provided for this phenomenon is usually in the context of 'interactive' (discussion oriented) knowledge networks where diversity in contribution reflects the diversity in the demographics of its members and of their different points of view.

In the context of topic pages, the most likely reason is that as a greater fraction of members contribute, the likelihood of contributors being from different geographical areas and different professional affiliations increases. These different groups have access to a greater pool of resources and this increases the overall level of resources contributed.

➢ There is evidence of topic/member maturity and learning effects as seen by the fact that there is an inverse link (corr. coeff. -0.52) between reject percentage and the number of resources in a topic and that the reject percentage seems to be decreasing over time.

An analysis of the findings support the critical mass hypothesis in both the topics analyzed – ICT and Aid effectiveness. The level of contribution depends to a significant extent on the contributions by both guides and the community in the previous period, indicating a presence of an avalanche effect in contributions. The subset of dependent variables that made the most significant contribution in some cases also included the contributions in time periods before the previous period (t-1), indicating the 'order' of the relationship might be higher than a month.

Further insights from this analysis include-
➢ In the ICT topic, *the independent variable seemed to be influenced more by the impact of contributions by the community (coefficient: 0.54, p value 0.07) than by the contributions by the guides*. This is important as it reflects the limit to which guide contributions can spur the members to contribute.

➢ One surprising result of the analysis was that in both the topics, *the number of people who participated in the previous period did not have any significant relationship to the contribution in a period*. This could probably be due to the fact that it is not easy for any member to find out how many different people contributed to the topic in a month.

4.7. Implications of findings and recommendations for restructuring the Topic Pages:

As mentioned earlier, the primary goal of the above exercise was to analyze the nature of contributions across topics across time and see if the characteristics of contributors, contributions and organizational processes play a part in determining the success of a topic- success here measured by the level and growth of contributions and membership across time. Based upon the above insights towards the following are the recommendations towards re-engineering the Development Gateway Topic Pages to achieve sustainability.

4.7.1.   Implications for Topic Rationalization:

The aim of the topic rationalization exercise is to see if the current categories of topics have spread the resources (links) too thin amongst themselves, and to see if it is worthwhile to group some topics together under a common umbrella to take advantage of scale effects.

The Development Gateway currently has thirty-five topics on a variety of subjects. These topics are characterized by varying participation levels in terms of membership, age and maturity of the topic and its members, the extent of community participation and the resources they contain.

As we saw in section 11.4.7, there seems to be some evidence of an avalanche effect in contributions, indicating that one of the many reasons for a member to contribute to a topic could be the presence of significant past activity within that topic. In other words, the marginal utility of a resource submitted to a topic that has not yet reached critical mass is bound to be lower than if the same resource were to be submitted to a Topic that has already reached critical mass or is on the way to reach it.

To illustrate using a simple example, suppose a Guide wants to submit a resource on say *"Information and Communication Technologies for Development in the Pacific",* (an actual resource submitted to the Development Gateway in Dec 2003), the guide has a choice of submitting it to either the 'Knowledge Economy' or 'ICT for Development' topic. In this case, the resource was submitted to the 'Knowledge Economy' topic which has one-thirds the total resources and one-half of the membership of ICT. The impact of this contribution towards eliciting member participation would probably be more if it were contributed to the 'ICT' topic.

Using the topic rationalization logic, it makes sense to group both these topics into one by subsuming 'Knowledge Economy' under ICT or creating a new topic. Similarly 'group able' topics are 'Business Environment' and 'Privatization', 'Foreign Direct Investment' and 'Trade and Development', 'e-governance' within 'Governance' etc.

Although the Topic Pages provide an option to submit the same resource to more than one topic the tendency of the member is to submit the resource to the topic where it is most relevant and it is to be expected that guides would submit resources to their own topics. This indicates the need to restructure the incentives for guides towards submitting resources to all topics that it might reference. This could be done either by measuring the number of postings outside their topic, or by assigning guides to groups of topics, where they would ensure posting of resources to all topics it is relevant to within that group.

Yet another way would be to use an ontology that takes into account the interrelationships across topics in a more meaningful way. To elaborate- the GSSD uses a proprietary ontology that maps out the domains of topics in sustainable development in an integrated structure that makes it easier to list a resource into multiple categories easily. The content submission process of the Topic Pages needs to provide for something like this.

To ensure proper grouping of topics, we would need data both on cross postings (how many resources on 'Knowledge Economy' are listed in 'ICT for Development') and on cross memberships (what fraction of members of 'ICT' are also members of 'Knowledge Economy'). This data is currently not available. However, a similar exercise on content mapping of the GSSD knowledge base (see section 5.2.1) suggested rich opportunities for rationalizing topics into higher-level groups.

In some cases, mature topics can even be split into sub-topics, where it is expected that each of the subtopics could reach critical mass by themselves. The important thing to remember here is that topics are not cast in stone, and rationalization can be used as an instrument of change to ·shape the characteristics of membership to drive overall network sustainability.

Topic rationalization should be an ongoing exercise and needs to be done taking into account the current characteristics of the topics and the way they would evolve. Topics that haven't reached critical mass yet might do so in the future, and it would be unwise to prematurely group them with other similar or related topics. Another factor that needs to be taken into account while grouping is the loss in content-specificity that is bound to arise. The challenge here is to group related topics to pool resources and members such that the higher-level topic does not become too general and mislead/confuse its members.

4.7.2.   Implications for community building:

The presence of an active, contributing community that promotes conversations and fosters relationships amongst its members is important for a knowledge network's success, both in terms of direct content creation and from the view point of incentivising other members of the community to contribute. Also as the avalanche effect indicates, it is important to display 'activity' levels to reassure members that their community is alive and thriving. An inactive and fractured community discourages its members from contributing and drives off new members. The Web site has to display changes in contributions each time the member logs on.

Studies mentioned in Section 3.1.1, 3.2.1 and 3.2.2 have indicated that the feeling of belonging to a community is one of the most important intrinsic motivating factors for contribution to a knowledge network. For the topic pages- this could mean a significant reduction in the burden of guides who have to constantly keep submitting new resources to their topics.

The first step towards community creation is to establish, maintain and grow the communication links amongst the members. However, most of the communication engaged in by the guides and the members are between themselves and the knowledge repository. There is very little member⇔member or member⇔guide interaction. Although tools for such interaction – like e-bulletin boards and provision for commenting on resources exist, they are not used very much. In fact, in the ICT topic, there have only been 105 discussions and 141 comments on topics till date- this in a topic with 6,000 members and over 6500 resources.

Although the topic pages provide for a variety of ways by which members could interact with each other and with the guides, there does not seem to be a definite feeling of a community as seen from online activity. A member survey of the reasons for non-participation might give us more insight into why this is happening. One possible reason for this could be that these tools are not easily visible, and many members may not know that these tools exist and it might be a good idea to display the latest discussions and comments on a topic in a more prominent area. Another possible reason is that community networking tools like Bulletin Boards and comments for documents typically face start up problems- they need to get to a certain level of activity before they become self sustaining. Creating these levels of activity is a challenge for any new network. One possible way to get over high start up costs are to 'engineer' crises within a topic – by posting a controversial point of view that spurs members into action leading to scale effects. Another way is to make guides and other institutional partners consider replying to discussions and posting comments on other member's resources as a part of their job.

Another reason for the low participation in discussion forums might be because of the high percentage of unregistered, browsing members. Since it is not possible to participate in discussions anonymously, valuable contributions might be lost. The easier it is to join a conversation, the more visitors will become contributors. The trade-off here is that junk postings might dominate. However this could be handled by introducing guide approval as an intermediate step. The larger issue here whether there is sufficient incentive for browsers to register. Registration should benefit the users. Websites like Slashdot enables its users to customize the comment display and write journals, and Advogato users rate each other on their perceived contributions to the cause of Free Software. Unless the Topic Pages make a compelling case for registration (currently the only benefits of registration are e-mail alerts and newsletters)- for e.g. a

personalized page each time the user logs on, where he can view only his topics, discussions and replies to comments, turning away possible anonymous contributions to discussions does not make sense.

Social capital theory (see section 3.1.1) indicates that another important motivation for people to contribute is the need for peer recognition. Creating an artificial hierarchy amongst members – even an elementary classification based on when they joined the community newbies, senior members, moderators etc., or a system of member ratings (or stars) based on number/quality of posts, often provide important stimulus for participation. The Topic Pages do not provide for any mechanism to distinguish between or rate its members.

4.7.3.   Implications for Resource Management:

Resource management here refers to the resources – effort and money – expended by the gateway on maintaining its topics. As mentioned earlier the Gateway has identified that "... *The cost of provision and quality assurance of Content information grows in relation to the amount of information provided, making it difficult to balance the need to provide more and better free information with the available funds."... "The Development Gateway needs to change the nature of this relationship."* In other words, as the network grows, the site ownership and maintenance activities of the Gateway have proved to be a bottleneck. The nature of the relationship has to change from an *owner-sponsored network* to a *community driven* one. The two recommendations mentioned above- Topic Rationalization and emphasis on community building could help in this transformation.

Topic Rationalization, by pruning and grouping the number of topics could concentrate the guide's resources on a few topics where their contributions could have the most impact. Lesser number of topics require fewer guides to scout the Internet for resources, and whose efforts could be diverted to the creation of 'value added' content or towards community building activities.

Community building as mentioned before reduces the burden of the guides. As the community matures, a significant fraction of the resources will be contributed by the community. The community's role is not limited to content provision alone and over time, the community could assume the moderation and validation functions currently performed by the guides. A meta-moderation system (like Slashdot) could be used to validate resources and moderate discussions and comments.

We will revisit some of the above observations and recommendations on the Development Gateway in chapter 7, where we compare the Gateway with another knowledge network with a different architecture- the MIT developed GSSD .The next chapter discusses GSSD in detail.

5. CHAPTER V: Knowledge Networks in Practice (II):The case of GSSD

5.1. Overview:

The Global System for Sustainable Development (http://gssd.mit.edu) is a knowledge networking and management system covering the domain of sustainable development, implemented over the Internet, and representing the work of GSSD partners all over the world.

Adopting a meta-networking strategy, GSSD provides networking facilities across stakeholder communities in order to help identify innovative approaches, enabling technologies, as well as new institutional, financial and regulatory mechanisms for meeting sustainability challenges that confront all countries[90].

At the core of GSSD is its knowledge base, consisting of a body of quality-controlled Internet resources on sustainability accompanied by abstracts of their content. These abstracts are organized within a consistent conceptual framework and translated into all supported GSSD languages (currently supported languages are English, Chinese and Arabic with soon to be included mirror sites for Japanese and French translations). The tags used for cross-referencing in the GSSD knowledge base allow efficient retrieval of these abstracts by user specifications and facilitate an understanding of the linkages among issues and problems, strategies, and solutions. In this manner, GSSD serves to enhance integration of alternative views of, and perspectives on, sustainability.

Surrounding the core knowledge base of GSSD is a set of processes and workflow –from content submission to translation to publishing. The translation process deserves special mention here. Any abstract submitted within GSSD is automatically routed through an inbuilt workflow to a translation partner, who translates the abstract into his own language. Over time, this results in a system where the content is available in all GSSD supported languages.

This process, along with the fact that GSSD is distributed over the World Wide Web (WWW) through a system of servers, called mirror sites, which allow users worldwide to select both the server location that provides them with the fastest access (bandwidth) and the language they most prefer, greatly reduces two of the most difficult barriers to knowledge, especially in developing countries.

---

[90] Source: The GSSD website: http://gssd.mit.edu

Operationally[91], GSSD is a coherent system consisting of integrated:

- Strategy for integrating and organizing knowledge related to the domain of sustainable development, in multi-dimensional, multi-sectoral, and international terms;

- Method to represent this knowledge by a plurality of interrelated concepts, and interrelationships that are organized in hierarchical form;

- Set of functionalities consisting of navigation and search mechanisms to operate over the system's quality-controlled knowledge base;

- Set of multi-lingual functions enabling non-English speaking users access to the same functionalities as English-speaking users.

- Modality of wide-area knowledge management, sharing and networking, across cultures, languages and disciplines;

- Mirror sites of GSSD located in different regions and/or countries worldwide; and

- Set of partnerships that are functionally specific and targeted oriented.

By building a decentralized network from existing sources of knowledge and experience, the GSSD helps users integrate untapped resources that increase local and global knowledge. Thus, the platform provided by GSSD is an efficient and effective tool to harness knowledge and resources that may reside locally, but are transformed into global assets for sustainable development.

5.1.1.   Information Architecture:

Computationally, GSSD follows two principles - to provide the user with an internally consistent logical procedure of obtaining access to information pertaining to sustainability and to develop a systematic method to facilitate that access. The GSSD architecture is structured as a hierarchical, nested set of relationships. The figure below shows the cross-sectional (multidimensional) view of the GSSD architecture for knowledge management.

GSSD maps out the domain of sustainable development into 14 Slices or core concepts that impact topics ranging from Industry, Agriculture, Energy, Water , Population, Conflict & Wars etc.

---

[91] N. Choucri, Mapping Sustainability: Toward a multidimensional content-based knowledge system to represent, track & deconstruct key domains and dimensions of 'sustainable development', last revised March 17, 2003. Source the GSSD Website: http:// gssd.mit.edu
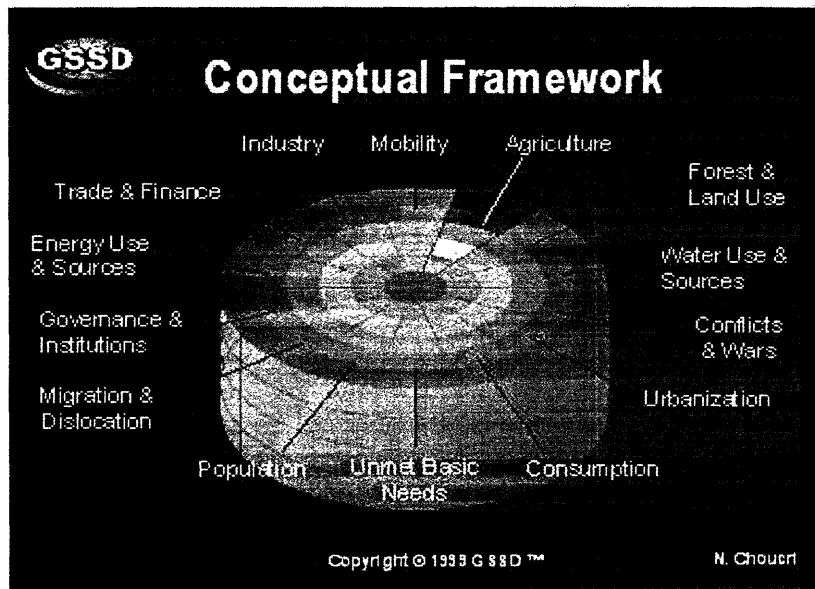
Figure 22: GSSD Conceptual Framework: Slices

The set of Core Concepts, listed above, is further differentiated into a set of dimensions or issue foci customized to the realities of each core concept. These dimensions are represented below as a set of concentric rings – 'Activities and Conditions', 'Sustainability Problems', 'Scientific and Technical Solutions' and 'Social, Economic, Political and Regulatory Solutions'.
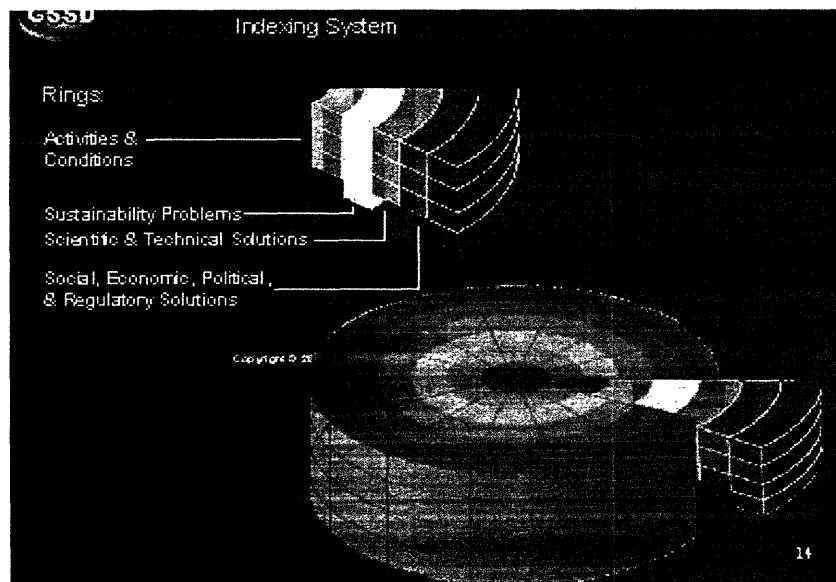
Figure 23: GSSD Conceptual Framework: Rings

The figure below represents distinctive knowledge-content or items at specific intersection of a Slice and a Ring (i.e. domain and dimension) as is shown for the slice 'Agriculture' and ring 'Sustainability Problems' below. This is also called a 'cell'.
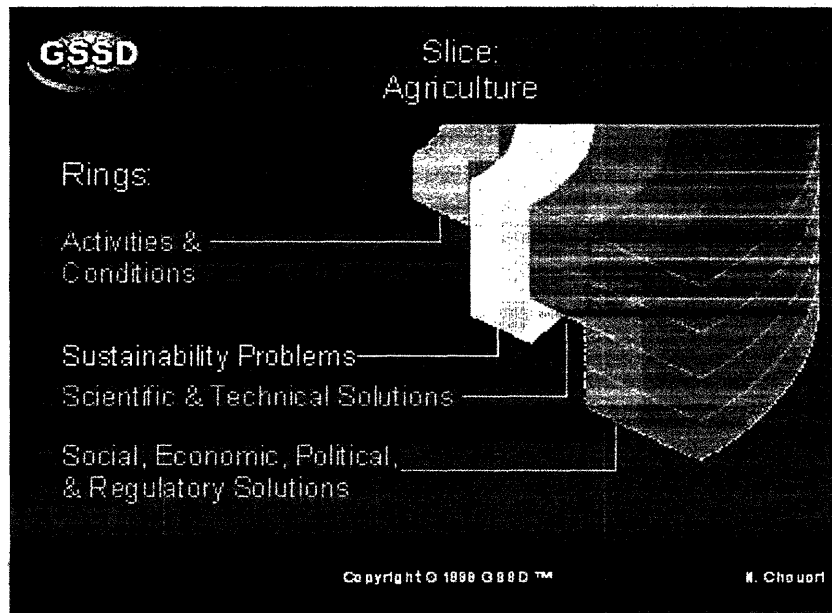
Figure 24: GSSD Conceptual Framework: Ring within Slice

GSSD permits drilling down further into a cell providing a further degree of granularity to the content through the mechanisms of a 'Concept' - referring to a specific item or issue within the Cell topic or a 'sub-concept'- referring to a specific element within the Concept level.
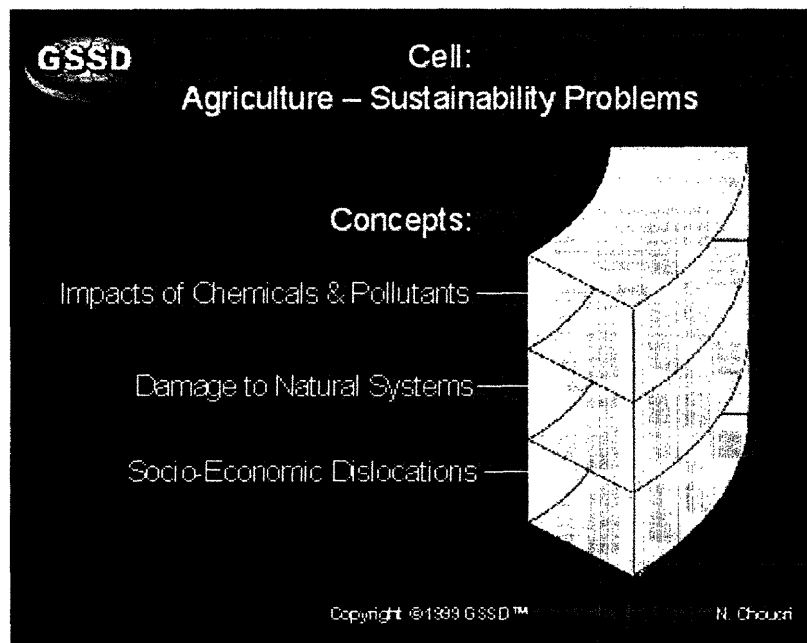
Figure 25: GSSD Conceptual Framework: Concepts
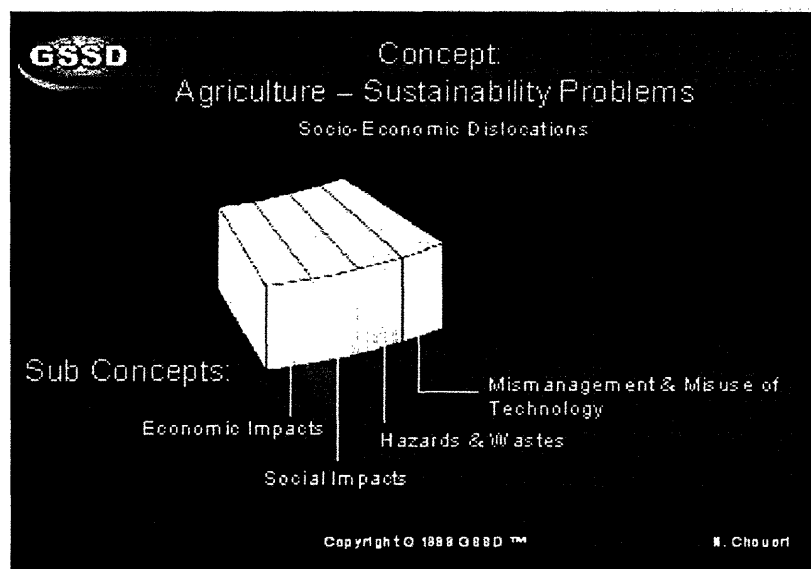
Figure 26: GSSD Conceptual Framework: Sub Concepts

5.1.2.   Users of the GSSD:

The different user types along with a brief description of their roles and the tasks they perform is within the GSSD System is summarized below. It would have to be noted here that very often a user plays multiple roles- for example editors could be translators and also submit content.

Role	Description	Regular Tasks	Software Used
REGULAR USERS	Who navigate the content in GSSD	Browse, search and submit abstracts	Web Browser
Content Editors	Manage all content submitted to GSSD.	Submit, modify, and review abstracts, and modify static pages.	Lotus Notes Client
Translator	Translate content on GSSD.	Translate English GSSD abstracts and static pages to other languages.	Lotus Notes Client
SYSTEM ADMINISTRATOR	Manage the day-to-day maintenance tasks associated with GSSD.	Network, server, database, group, and user maintenance.	Lotus Notes Administrator
Designer	Design databases and web pages.	Creating new, and updating existing, design templates.	Lotus Domino Global Workbench, Designer

Table 6: Users types and roles within GSSD

111

5.1.3. Technology and Network Architecture:

GSSD has hub and spoke architecture, with the central hub server (called GSSD Stage) responsible for keeping the content in each of the mirror sites in synch and current. The rationale behind this architecture is to meet one of the primary goals of the system- to have a copy of GSSD in *all* languages in *each* of the locations for easier access. Thus a GSSD mirror site in China would contain all language databases (English, French, Chinese, Japanese and Arabic currently). We shall revisit the architectural choice of GSSD at a later stage in this report and discuss if the assumptions under which this architecture was chosen still holds.

The GSSD has partnered with Lotus Development Corporation (now a division of IBM) for the development of the application. One of the driving forces behind using the Lotus platform for GSSD was its localization capabilities, where localization refers to the process of creating a multilingual application. The Lotus Domino Global Workbench software allows for the internationalization of the GSSD knowledge network by providing GSSD's knowledge base in multiple languages.
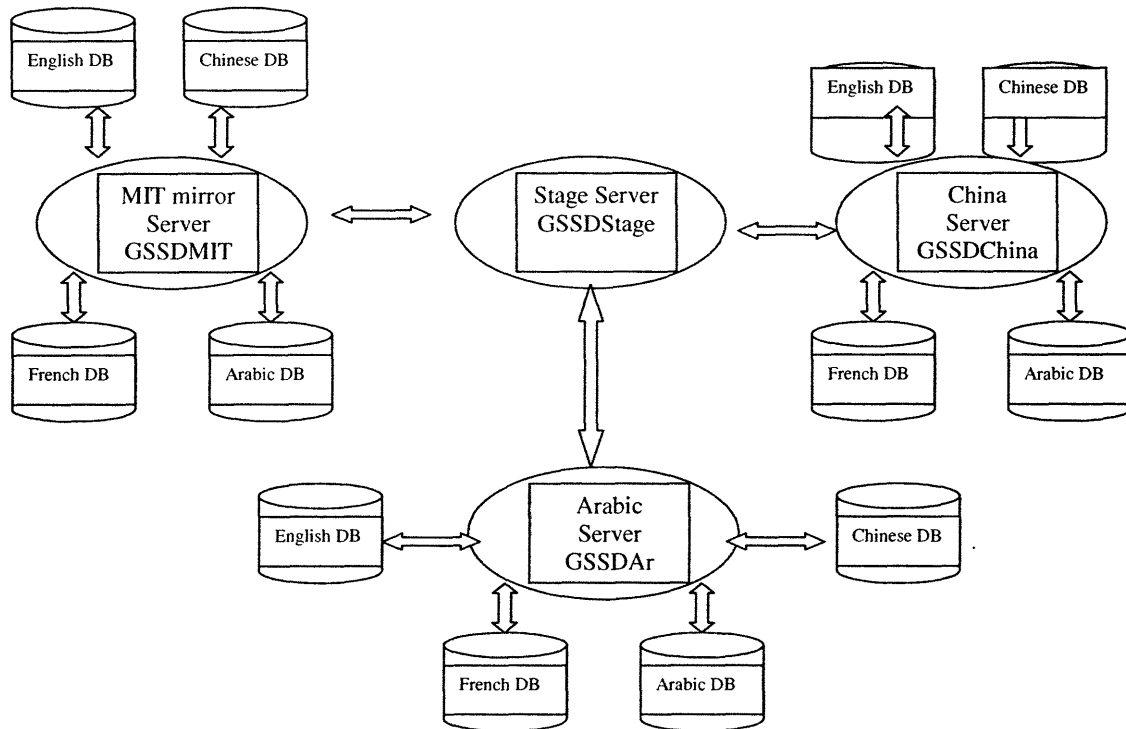
Figure 27: GSSD: Technology and Network Architecture

Among the many reasons for using the Lotus platform for localization, as opposed to just going into an application and translating the content by hand, includes the need for specificity of skills. The Domino Global Workbench provides the facility for creating and maintaining a multilingual glossary of terms, definitions and associated design elements, thus allowing a developer to create a new language mirror site very easily. The Workbench standardizes the terminologies related to sustainable development, reducing the chances of misunderstanding and helping translators translate their content much faster.

Another important reason for choosing Lotus Domino is its excellent content replication and synchronization features. The Domino server ensures that the content on all mirror sites is replicated (an abstract added in one mirror server gets added to all mirror sites on replication), and synchronized (a change made in one language in one mirror server gets replicated over all servers into *all* other languages). These powerful features ensure that the content within the network is both current, and consistent, and coupled with a strong document management platform, makes for easy content management with very little administrative overhead.

The third reason for the choice of Lotus Domino was its workflow and collaboration capabilities from document submission to revision and translation to publishing. Lotus/IBM is amongst the leaders in the knowledge management/collaboration tools.

5.1.4. Organization of GSSD:

The GSSD Consortium is a loosely managed MIT led initiative motivated initially by a set of actors, including Governmental Institutions (consisting of select sponsors of the MIT Symposia on Global Accords for Sustainable Development), Business & Industry Partners (IBM, XEROX-PARC, etc.) and Research Institutions. Jointly they are responsible for the operational strategy and implementation of GSSD, as well as the development and deployment of its knowledge assets.

At the more operational level, current partnerships of GSSD below are:

GSSD Partnership Type	Description of Partnership	Example of Partner[92]
Content Partners	Combs the Internet for appropriate content for the GSSD Knowledge Base, then abstracts, indexes, and publishes it to the web. Also reviews submissions from web users. Content Partners may also be an office in an organization responsible for updating their own content on GSSD.	American University, Beirut, Lebanon
Translation Partners	Translates new and modified materials in both the GSSD interface and it's content.	Écoles Des Mines, St. Étienne, France
Mirror Site Partners	Maintains a server, live to the web and integrated into the GSSD system, which contains a replica of the GSSD system. These replicas are automatically updated daily.	Ministry of Science & Technology, Beijing, China
Development Collaborators	Provides development support, in the technology, experience, collaboration, etc.	Lotus Corporation, USA
General Supporters	Provides general support in terms of Participation ('equity sweat'), in-kind support, or direct financial contribution	Baker&McKenzie AT&T Xerox Corp.

Figure 28: Organization of GSSD

Thus, organizationally, GSSD comprises of a network of different kinds of partnerships managed by a core group within MIT, with different kinds of institutional support – from educational institutions to governmental support (The China Ministry of Science & Technology is responsible

114

for the Chinese Mirror Site operations). The core day to day group within MIT comprises 1 full time Faculty member, 1 full time employee and two student researchers and is supported by the MIT Dept of Political Science's technical staff as well as UROP students for facilitating content provision.

5.2. Content and Traffic Analysis:

5.2.1. Content Analysis

Objective: To analyze existing data within the GSSD system with a view to measuring the effectiveness of its ontology. Effectiveness defined here as the ability of the ontology to help the user find the content he wants.

Problem Definition: The GSSD is a web based multilingual and multisite platform through which users from all over the world can submit 'abstracts' of websites on Sustainable Development for inclusion in the GSSD Database. These submitted abstracts are evaluated by the GSSD team, modified if necessary and catalogued appropriately using the GSSD ontology. Currently in its fifth year of existence, the GSSD has published over 2000 abstracts on a wide range of topics related to sustainable development and can be considered to be an indication of the state of the literature existing on sustainable development on the WWW.

Revisiting the operating principles of GSSD (as outlined in the preceding section), the purpose of the GSSD is to provide a coherent

- strategy for integrating and organizing knowledge related to the domain of sustainable development, in multi-dimensional, multi-sectoral, and international terms;

- method to represent this knowledge by a plurality of interrelated concepts, and interrelationships that are organized in hierarchical form;

This exercise is an attempt to gauge the effectiveness of the ontology with an aim to see if the above goals have been realized and whether the current classification system needs to be modified. This approach could also be viewed as a generalized method towards analyzing other similar classification systems –such as the World Bank's Development Gateway's Topic Pages and the UNEP's Grid System on sustainable development. This approach also could be a precursor to developing 'dynamic' ontologies- ontologies that adapt continuously based on the content that flows through them.

Analysis Methodology and Approach: The GSSD Lotus Notes Database, was exported and processed to yield a text file with each row of the file representing one abstract. Each abstract could reference one of more 'slices' and one or more 'rings'. The slices being the domains of sustainable development('Agriculture', 'Industry' etc.) and the rings being the dimensions ('Activities & Conditions', 'Scientific & Technical Solutions' etc.) .On each pairwise combination of slices and rings, the number of abstracts referencing both were calculated, and the results summarized in a table. The result of this exercise helps us visualize the extent of interrelationships in the knowledge base content.

Results: The sample table for the Slices is represented below-

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	Total Abstracts
A	7%	5%	9%	6%	11%	6%	7%	8%	5%	6%	7%	8%	7%	10%	904
B	3%	26%	3%	2%	3%	4%	1%	9%	2%	4%	3%	5%	3%	3%	453
C	7%	5%	5%	6%	7%	8%	6%	6%	5%	5%	5%	6%	8%	8%	764
D	11%	6%	13%	13%	11%	10%	24%	7%	25%	6%	10%	9%	8%	10%	1701
E	11%	6%	8%	6%	11%	7%	7%	8%	6%	6%	6%	8%	7%	11%	934
F	5%	6%	6%	4%	5%	4%	4%	7%	3%	6%	6%	8%	8%	6%	650
G	13%	6%	13%	23%	12%	11%	9%	7%	26%	6%	11%	8%	8%	12%	1689
H	3%	6%	2%	1%	3%	4%	1%	5%	1%	4%	3%	4%	4%	3%	341
I	8%	5%	9%	20%	8%	7%	20%	6%	7%	5%	8%	6%	6%	8%	1347
J	5%	6%	5%	3%	5%	7%	3%	8%	3%	26%	5%	9%	8%	4%	747
K	5%	5%	5%	4%	5%	7%	5%	6%	4%	4%	17%	7%	6%	5%	726
L	7%	7%	6%	4%	6%	9%	3%	9%	3%	9%	7%	6%	9%	6%	720
M	6%	5%	7%	3%	5%	8%	3%	8%	3%	7%	6%	8%	11%	6%	705
N	10%	6%	8%	5%	10%	8%	6%	7%	5%	5%	6%	7%	7%	7%	839

Where -

Agriculture	A	Migration and Dislocation	H
Conflicts and War	B	Mobility	I
Consumption	C	Population	J
Energy Uses and Sources	D	Trade & Finance	K
Forest and Land use	E	Urbanization	L
Governance & Institutions	F	Unmet Basic Needs	M
Industry	G	Water Use & Sources	N

Table 7: Pairwise comparison of Slices

The above table should be read vertically (by columns). For example, the entry in Row C, Column D (6%); should be interpreted as, of all the 764 abstracts in the database that contain

the slice 'Energy Use and Sources', 6% of them also contain the slice 'Consumption'. The table is asymmetric - the entries for Row C, Column D are different from those of Row D, Column C . This is because, the figure 13% represents the percentage of abstracts that contain 'Energy Use and Sources' as a percentage of abstracts that contain the slice 'Consumption' and is higher than the 6% figure from above.

Interpretation of results and recommendations towards re-engineering the ontology:

For researchers trying to map a new domain, one measure of the effectiveness of their ontology could be the distribution of results in the above table. A more equalized distribution could indicate a more effective classification, where as one with 'lumps' could indicate a high dependency across two slices.

Lumpiness could be due to a natural convergence between two domains (such as say energy and environment) and sometimes afford a chance to make the slices more granular (say dividing 'energy' into 'power generation' and 'energy sources').

Also as the above example illustrating the difference between (C,D) and (D,C) indicates, one could gain insight into the 'directionality' of the slices, with one slice being more important to another in a pair-wise relationship.

Both granularity and directionality have important implications towards Topic Rationalization and indirectly the user navigation experience. For example the above table could be useful in creating a navigational structure for researchers from one domain wishing to see how many, and what type of content links to other slices and domains.

	A	B	C	D	Total Abstracts
A	56%	16%	8%	10%	587
B	5%	59%	5%	6%	313
C	27%	15%	44%	40%	980
D	12%	16%	35%	46%	860

Table 8: Pairwise comparison of rings

Where -

Activities & Conditions **A**
Sustainability Problems **B**
Scientific & Technical Solutions **C**
Social/Economic/Political & Regulatory Solutions **D**

117

The above table is the result of a pairwise comparison of rings. The interpretation of this table is similar to the one we saw before for the pairwise comparison of slices. For example, the entry in row B, column A can be read as " Of the 313 abstracts submitted to GSSD on sustainability problems, only 5% of those also mention the activities and conditions that led to those problems".

The above table also provides one measure of the classification efficiency of the GSSD ontology. As the size of the entries along the diagonal of the above table indicates, the GSSD knowledge base contains a high fraction of abstracts that fall within the categories within the ontology and in this sense, the GSSD ontology is highly focused.

A further area of research could be in linking the above table to each slice of table I as this could provide useful insight into the knowledge space characterizing a domain. For example, it might lead to answers to questions on the extent of solutions available to address problems within a domain, the emergence of new problems within a domain and the change in the knowledge space within the domain over time etc.

GSSD's competitive advantage relative to other similar meta-networks on sustainable development is its powerful ontology and classification system. Continuously updating the results of such an analysis into revisions of the ontology is needed if GSSD expects to sustain this advantage. The results of the rudimentary content analysis above is the first step in this direction. Further work towards content analysis to further fine tune the ontology could include-

- Formal Cluster analysis to determine size of optimal clusters to characterize data :The current cluster analysis is done on pair-wise comparisons of abstracts. The content analysis will have to consider more dimensions in its comparisons- as each abstract can reference more than one slice. Sophisticated graphical analysis tools for content affinity analysis need to be used for this purpose.
- Analyzing Rings and Slices together: The current content analysis looks at Rings and Slices in isolation. By combining Rings and Slices together as a unit of analysis, one could get answers to questions like 'what fraction of abstracts linked to both energy and industry are also technical and scientific solutions vs. those which are social/economic/policy solutions'.
- Characteristics across other dimensions like location/time etc. The content analysis described above ignores the characteristics of the user submitting the data. It would be interesting to know the answers to questions like 'are developing countries contributing more than developed nations to certain domains?', 'which countries lead in providing technical solutions/ policy solutions'? etc.

118

5.2.2.　Traffic Analysis and Network Mapping

Objective:

I.　to analyze web traffic patterns and user profile to the GSSD website and

II.　to determine the relative position of GSSD with respect to other knowledge sources on sustainable development on the Internet.

Traffic Analysis results and interpretation:

Server log file data for a 410 day period from 1st Jan 2002 – 15th Feb 2003 was obtained and analyzed using a commercial log file analysis tool. The summary of results is provided below-

		Description
First hit	12/31/2001 12:00:46 AM	Time of first hit
Last hit	02/15/2003 05:27:41 PM	Time of last hit
Hits	421032	Total number of objects accessed on web site
Page views	279209	Total number of page views
Visits	64659	Total number of visits (continuous page views)
Page views per visit	4.32	Average number of pages viewed per visit
Time per visit	0:02:51	Average duration of a visit
Visitors	13382	Total number of visitors
Visits per visitor	4.83	Average number of visits per visitor
One-time visitors	11040 (82%)	Visitors visiting only once
Time spent per repeat visitor	0:13:48	Average total visiting time per repeat visitor
Visits per day	156.94	Average number of visits per day

Table 9: GSSD Traffic Summary 2002

The above table indicates that on an average day the GSSD gets around 160 visitors a day, a healthy number considering GSSD does not aggressively promote itself, the content is intellectually 'heavy', and relies largely on pull (through branding, reputation and a satisfying experience) marketing for its visitor base.

However, the above figures also indicate a troubling statistics- the number of one-time visitors is a very high 82%, which means that 6 out of 7 visitors who visit GSSD never come back. What is interesting is that the number of visits per visitor is 4.8, which means that even with a 82% one time visitor percentage, the number of visits by repeat visitors is so large that the average comes out to 5 visits/visitor. This proves the existence of two very different classes of users- browsing ·

119

members who are usually one time visitors, and 'heavy', repeat members- constituting 1/7th of the GSSD visitor base, but who use the website heavily. An even more encouraging statistic is the amount of time repeat visitors spend on the site- a whopping 13 minutes.

A sustainable knowledge network needs a high percentage of repeat visitors whose sustained participation is necessary for a critical mass of members and resources to build up *and* a steady supply of new members to compensate for natural attrition and who go on to become repeat visitors. The high browsing/repeat visitor ratio is worrying in terms of recruiting new members into the network, and GSSD has to look for the reasons why this is happening. The reasons for this phenomenon could be either/all of marketing factors (attracting the wrong kind of audience in the first place), bad user experience (not providing what the visitors are looking for, or cumbersome navigation and search) technological factors (slow or unreliable website because of downtime), content/knowledge network factors (lack of relevant content or not enough content).

It is possible to find some partial answers to the reasons why this is happening. An analysis of the most frequent exit pages (the pages last seen by the visitor before he exits the site) yields the following results.

Figure 29: Most frequent exit pages
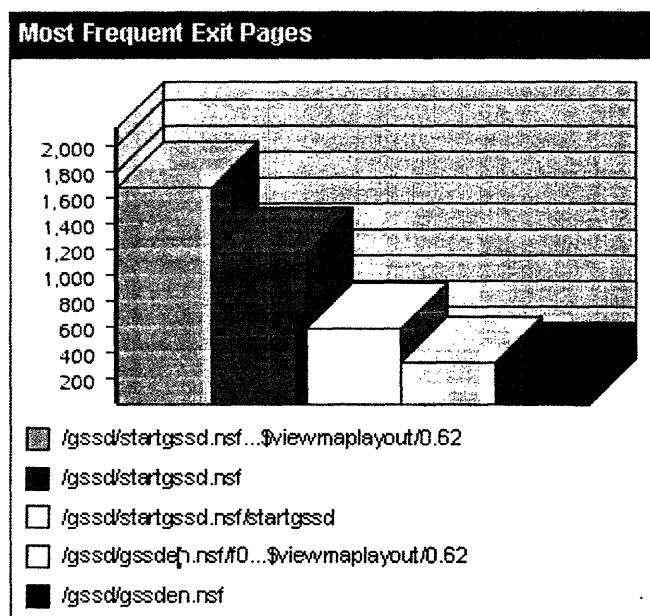
It can be seen that the most frequent exit pages all occur within the first two navigation levels of GSSD – i.e within the introduction page(where the visitor has to choose from amongst a list of languages and mirror site locations) and the start page. The fact that most visitors exit from these pages indicates that the user who visits the GSSD website either faces confusion (over the

120

choices), or frustration (over the fact that he has to go through multiple levels) to get the information he wants.

This is not a serious problem for repeat visitors who are used to the interface, but this might be a significant deterrent to attracting new visitors to come back again. Clearly, GSSD will have to look at alternative design choices in the first two pages to improve new user experience.

Network Mapping Analysis:

The aim of this exercise is to find the relative position of GSSD vis a vis existing content environment on sustainable development currently existing on the internet. This approach maps out the domain of knowledge on this topic on the web and provides one with a topology of similar content which could form the basis of collaboration and reverse linking, competitive analysis and benchmarking.

It could provide also reasons as to the reasons for the high number of one time users could be due to the nature by which individuals come on to the GSSD network in the first place. Specifically, we seek to find out the possible approach paths to the GSSD website (either by directly typing http://gssd.mit.edu, or by redirection from another website, or through a search engine). The log file data for redirection and direct entry into GSSD is not available, but a look at the search engine keywords used to come to the GSSD website gives us some insights. The below is the analysis on the keyword data for a period of two months (September- October 2003)
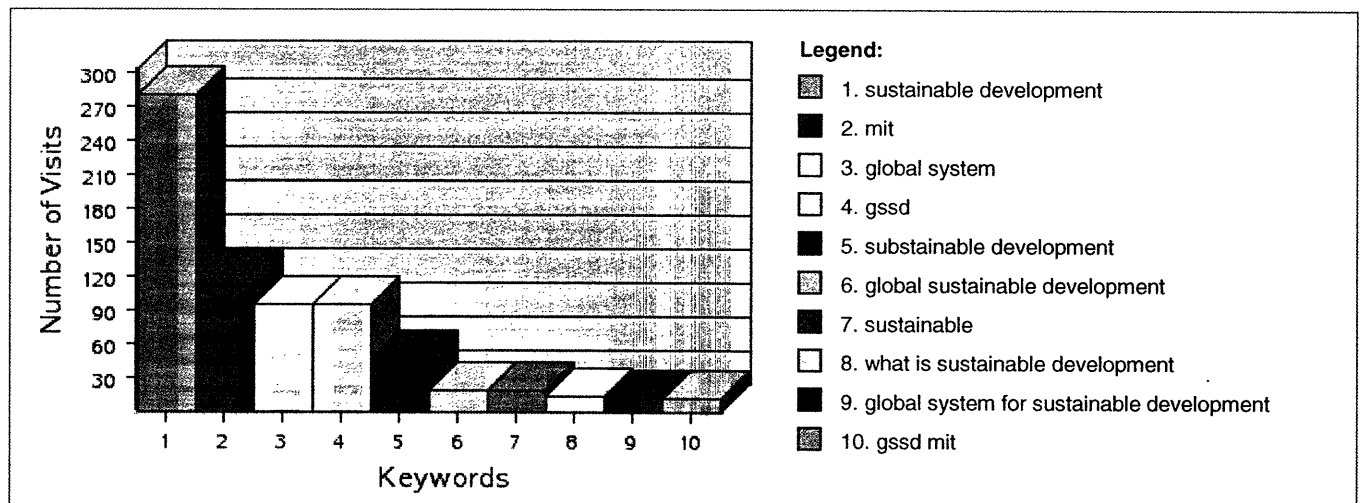

Figure 30: Most frequently used keywords to find GSSD (1)

	Keywords	Visits
1.	sustainable development	280 (4.9%)
2.	Mit	118 (2.0%)
3.	global system	94 (1.6%)
4.	Gssd	94 (1.6%)
5.	substainable development	46 (0.8%)
6.	global sustainable development	19 (0.3%)
7.	Sustainable	19 (0.3%)
8.	what is sustainable development	14 (0.2%)
9.	global system for sustainable development	12 (0.2%)
10.	gssd mit	11 (0.2%)

Figure 31: Most frequently used keywords to find GSSD (2)

We see that users who come to GSSD come in using either a combination of general keywords (sustainable development, sustainable etc.), or very specific keywords (GSSD, MIT etc.) What is curious though is that none of the popular keywords reference any of the content within GSSD. This is usually the largest and most important source of traffic for most knowledge networks (most knowledge network resources are indexed and made available to search engine robots). GSSD has only recently (since October 2003) opened its database for public indexing. This is expected to yield a substantial rise in user traffic.

A visibility analysis which computes the number of links from other sites pointing to GSSD (homepage and US server only), the number of pages returned for the exact phrase (i.e. pages containing the expression- within and outside GSSD pages) and the page rank of key terms is reproduced below for six popular search engines. This analysis was performed on the 28th of October 2003[93].

		google	**fast**	**ixquick**	**altavista**	**yahoo**	**msn**
Links to GSSD Main Mirror	link:gssd.mit.edu	160	474	304	332	160	n/a
	Link :gssd.mit.edu/GSSD/gssden.nsf	0	21	13	20	n/a	n/a
References	"global system for sustainable development"	895	574	433	399	748	409
	"sustainable development" (directories: refs, rank)	5, 19'	n/a	n/a	2	1, 17'	n/a
Page Rank	intitle:"sustainable development"	33'	15'	n/a	57'	30'	30'

Table 10:GSSD search engine positioning

---

[93] Vincent Maugis (2003): Web-based Content, Behavior and Context Mapping and some Comparisons with Matrix-based Analysis, Working Paper.

Network Mapping analysis, by using a cartographic tool, helps map out where GSSD stands with respect to other providers on sustainability resources on the web. The following figure is a result of one such analysis-
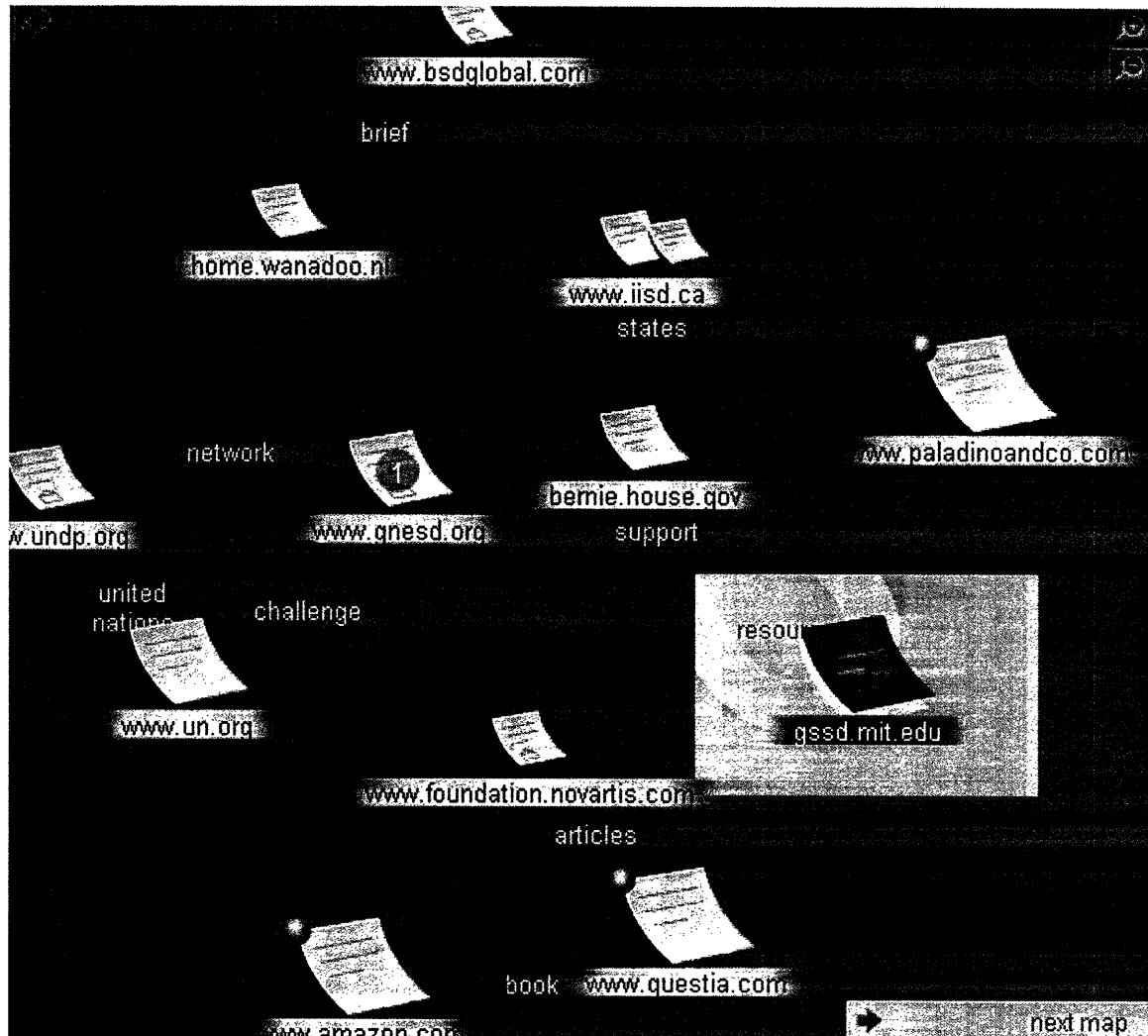
Figure 32: GSSD's position in the sustainable development domain: Kartoo Output

The figure above is the output from Kartoo – a web based cartographic search engine available at http://www.kartoo.com .It indicates where GSSD stands with respect to others on the search phrase 'global sustainable development'. Based on this map, one can clearly see that while GSSD is geared towards users seeking 'resources' on sustainable development, the UNDP and the UN websites are for networking and for describing the challenges respectively. Amazon and Questia provide information on books related to sustainable development.

The results of the visibility analysis and the network mapping exercise helps the network manager plan corrective action to increase the results of his website for the key words most important to

123

him on the most popular search engines. Moreover, by examining the link to(the list of sites which link to the network the most), the network manager could identify co-promotion opportunities that could provide a steady source of traffic in a more sustainable way.

5.3. Implications of results towards re-engineering the GSSD network:

The following is a summary of insights from the traffic analysis and the network mapping exercise on the GSSD database outlined in the previous section:

> Although overall traffic to the website is healthy the number of repeat visitors is low, as compared to similar statistics for other networks (Development Gateway has a repeat visitor percentage of up to 30%)

> The repeat visitors contribute heavily to traffic (visits and page views) but contributions to the resource pool are low (although this might have been due to technical glitches- which were corrected in Aug2003- the level of provision is still low). Also the amount of time repeat visitors spend on the site and the number of pages they view indicate that free riding is a major problem. Resource provision is now almost wholly by GSSD partners and this trend could be reversed.

> GSSD rates highly on sustainable development on all major search engines and has good comparative recognition w.r.t. similar networks that provide resources on sustainable development. The network should capitalize on this opportunity to convert new members to repeat visitors.

The problem facing GSSD is similar to the problem facing the Development Gateway. Both models of knowledge networking rely heavily upon actors within institutions for providing content. The Gateway has explicitly recognized this in their business plan and they admit that their current model will not scale if they wish to grow. One of the contributions of this thesis was to show a way in which the community participation in the Gateway could be increased to assume a part of the burden of knowledge provision.

GSSD faces a similar problem of under provision of knowledge resources due to a lack of community participation. However, the important difference is that the kind of resources submitted into GSSD are more selective, the process of approval more rigorous and the GSSD user much more discerning than the average Gateway user. A more fundamental difference lies in the nature of knowledge within each of these networks. While the knowledge within GSSD is more akin to

124

the 'basic research', the knowledge the Gateway collects and disseminates is more akin to 'applied research'.

To explain, the GSSD's core competence is the value added to the content through its powerful ontology. The users of GSSD use the ontology to navigate and search for content and appreciate the results which not only provide an abstract or meta data for the content, but also the ability to view the linkages and interconnectedness between the domain they are interested in and the other domains it may impact and the ability to view the same content in more than one language. Moreover, the GSSD is selective about the data it includes into its knowledge base- news reports and press releases – acceptable to the Development Gateway are not allowed within the GSSD knowledge base.

Clearly, the GSSD is aiming for quality over quantity, and moving to a community model of resource provision raises fears that the quality might be affected. However, we are still faced with the problem of continually populating the knowledge in a scalable manner. It must be realized that attracting a larger audience will not lead to a deterioration of the quality of content as the rigorous multi-stage approval process will still be used to weed out unworthy submissions. It must be recognized that GSSD only stands to gain irrespective of the quality of the users it attracts(as long as it does not turn off 'serious' users). GSSD should clearly 'hook' its users by attracting as large an audience as possible and 'bait' them with its features- a comprehensive, intelligent, quality controlled knowledge base.

Thus the key challenge within GSSD is to attract more visitors and to retain them. Attracting new visitors can be done by continuing to 'market' the knowledge base on search engines, implement link exchange programs and by e-mail marketing. However retaining visitors and making them visit again is a much harder task. The following is a list of recommendations towards increasing the number of repeat visitors the GSSD knowledge base.

5.3.1.   Solutions towards community building:

This category of recommendations deal with the issue of creating a sense of community. The logic behind this category of solutions is repeat visits would lead to increased contributions, and that to encourage repeat visits the 'stickiness' of the network should be increased. One way of increasing stickiness is to build a feeling of community. Within the GSSD the barriers to stickiness that have been identified are-

➢ **The lack of persistent identity:** Communities are created when individuals build and maintain relationships within a networks. Researchers[94] have identified that interpersonal trust is a central characteristic of relationships that promotes effective knowledge creation and sharing in networks. However to create a feeling of trust, one needs to have a unique identity within the network. Currently GSSD does not recognize unique identities externally – it is not possible to find out who submitted a resource, nor is there a way to log on to the network and be recognized by the 'system'. The first step towards building a community would involve a registration mechanism where users register their preferences and profile with the network with an option to disclose their profile to all other registered users.

➢ **The lack of connections:** There are no mechanisms for social connections within GSSD. Interactions are one way- between the user and the knowledge base. New connections could be formed by allowing for comments on submitted resources or by launching new discussion groups.

➢ **Lack of Activity indicators:** As described previously in the Gateway case study, indicating high levels of activity could spur new members to action and nudge the network towards critical mass. Currently, the GSSD knowledge base does not provide any mechanism to determine the level of activity since the last visit. There should be a provision for displaying the most recent resources added, the number of resources in each of the domains and the number of new members added since last log on, in a prominent location.

The second category of solutions deal with changing the organizational processes for improving community participation.

5.3.2. Re-engineering work processes

Work processes here refer to the content submission policy, the content review process, the process for translation and publishing. To a certain extent it also refers to the user search and navigation process. The two recommendations in this context are

➢ **Redesigned Layout:** The current layout of GSSD involves 6 mouse clicks before the user can get to the 'search GSSD' page. This presents 6 opportunities for the user to leave the site. An analysis of the top exit pages have shown that most people leave the website after the first two pages. Clearly to stop this, the search feature should be made available in the first or second page.

---

[94] Tsai, W., & Ghoshal, S. (1998). Social capital and value creation: The role of intrafirm networks. Academy of Management Journal, 41: 464-476.

➢ **Content Submission Process:** The current content submission process involves filling a form with over 15 fields- a burdensome process for any user. This form should be simplified by decreasing the number of fields. This might increase the burden on the reviewed. The trade off here is between increasing the quantity of submissions into the knowledge base vs. increasing the work load on the reviewer. Clearly a balance has to be struck.

5.3.3. Re-engineering Partner Incentives:

The third category of solutions involves re-engineering the incentives and motivations of the network's institutional partners and their representatives. The current motivation for the translation/editorial partners for collaboration with GSSD are primarily due to a combination of factors ranging from a genuine desire to participate in an initiative that provides a holistic view of knowledge on sustainable development, a desire for peer recognition, a personal interest in the success and popularity of the knowledge base, a desire to promote their own language or region (digital inclusion), existing institutional (and in some cases governmental) arrangements and a desire to be affiliated to the MIT/GSSD brand name.

Of all the motivating factors, the most important ones are those where the actors genuinely believe in the power of GSSD to provide a new way for representing the state of the domain of sustainable development and where they have internalized the values that GSSD stands for (an integrated perspective that cuts across domains, geography, languages). As we have seen before commitment by internalization is the preferred way of cultivating commitment behavior as the synchronization of values are more long lasting than hard rewards, punishment or social recognition and ensure that the knowledge worker is invested at the deepest level in the prescribed behavior. In the context of GSSD this ensures that the knowledge worker will try to maximize one's value added contributions and will not artificially inflate the quantity of contributions or skimp on the quality.

It is important to recognize that all of these are 'soft' incentives and the temporal lag between the motivation and the actual deliverance of the result could pose problems for ensuring sustained participation in the network. Soft incentives take time to deliver and need to be continuously reinforced for them to take effect. GSSD has been effective in rewarding genuinely committed partners by inviting partners to conferences within and outside MIT. However, more needs to be done. Other ways of engaging committed partners could be by initiating institutional joint projects on similar topics, providing access to source of funding, increased networking options for partners where they interact with other MIT faculty/researchers, providing access to scarce resources etc.

127

Thus far we have seen that a community oriented participation model coupled with a strong institutional network backbone of partners can provide GSSD with a sustainable model. For this to happen, there needs to be a strong top management commitment, a long term vision for the network that clearly states objectives, timelines and resources needed and a funding/sponsorship plan in place. The next chapter looks at the learnings from the Development Gateway and GSSD cases and attempts to develop an integrated framework for re-engineering knowledge networks to attain sustainability.

6. CHAPTER VI: An integrated framework for re-engineering e-knowledge networks

The preceding two chapters discussed the specific cases of the Development Gateway and the GSSD. This chapter builds upon the results of the analysis performed in the two cases and attempts to compare and contrast the different models towards sustainability they have each adopted. The results of this comparison and the learnings from the previous chapters are then incorporated to provide an integrated framework for evaluating the progress of a knowledge network towards sustainability.

6.1. Comparing GSSD vs. Development Gateway

Although both the knowledge networks described above -the Development Gateway (DG) – both Topic Pages and the Country Gateways **(as their decentralized architecture, nature of partnerships and the mirror site network is similar to that of GSSD's)**, and the GSSD, aim to provide an interactive portal for information and knowledge sharing on sustainable development, there are important differences in their objectives, strategies and operating processes and institutional configurations. The goal of this exercise is to compare and contrast their approaches along the parameters of Overall Objectives & Strategy, their Information Architecture, the Nature of institutional partnerships, the Technological Choices they have made and their Organizational Structure.

6.1.1. Overall Objectives & Strategy:

The primary difference between the two knowledge networks is in their approach to knowledge itself, which indirectly stems from the differences between MIT and World Bank as institutions and their missions are reflected in their respective system.

MIT's GSSD is viewed as a service mainly for researchers in the sustainable development community.[although notable exceptions exist - GSSD has been used by OECD-IEA as framework to organize its technology agreements; by the Government of Bolivia to put in place the Bolivian GSSD and its interactions with the Chinese government] - reflecting MIT s central mission as a knowledge-based institution. On the other hand, the Development Gateway views itself as offering a hands-on application to international policy makers, project managers, local government authorities, private sector representatives, and researchers amongst others. MIT

clearly states that is not in the direct technical assistance business, but through its research mechanisms uses GSSD as a capacity building mechanism by way of creating new development strategies. Therefore, although the DG has a broader set of features- specifically apart from features facilitating the exchange ideas and knowledge, it also provides features that allows one to find development projects and explore business opportunities (by providing a global online marketplace), the GSSD is more focused on efficiently representing content and has differentiated itself through its ontology and its unique way of mapping the complex domain of sustainability.

A stakeholder analysis of the Development Gateway initiative gives an indicator of the key interest groups they intend to serve and the service they provide

Governments	
Need • Insufficient information and tools to improve government processes	*Value Propositions* • Platform for e-procurement and other e-government applications • Information on best practices on governance and administration for governments
Donors	
Need • Uncoordinated, overlapping, and sometimes ineffective aid	*Value Propositions* • Transparency and coordination of development information and activities through a multi-agency project database
Private Sector	
Need	*Value Propositions*

Untapped commercial opportunities for the private sector involvement in development projects	Expand access to business opportunities Reduce the transaction and search cost in doing international business for exporters and investors Improve access to tools for e-commerce

Table 11: Stakeholder Analysis

The very different markets the two initiative service influences their overall strategy in terms of the marketing plan, their plans to source funds and the technology strategy and architecture employed.

6.1.2.   Comparison of Information Processes:

The information architecture of both initiatives is heavily dependent on (i) the classification systems and ontology, (ii) the content and editorial policies currently in place and the control wielded over the content published (centralized vs. decentralized etc.) (iii) the nature of partnerships, (iv.) and the technological choices adopted. We will attempt to compare and contrast the way in which each the architecture of both systems has evolved with respect to these parameters.

Classification/Ontology:

The GSSD proprietary classification system has been described earlier in some detail. The content structure of Gateway is based on 35 topics and 4 crosscutting themes generally recognized as development priorities. Content can be further classified in to sub-topics. Items can also be accessed by type ('Resource categories'): 'Data and Statistics', 'Documents and Reports', 'Events and Discussion', 'Forums', 'Get Involved', 'How to / Tools', 'Organizations, Networks, People', 'Programs and Projects' and 'Publications and Multimedia'. For each content item a summary is entered, and users can access details on the contributor to the item, as well as registered users' comments on the resource.

The Development Gateway's ontology is however largely based on the nature of the content, and it ignores the meaning and context. In this sense, the ontology is under developed as compared to GSSD's ontology which is a powerful way to look at broad themes within a document that cut across many 'slices' in the domain of sustainable development. Given the nature of the ontology of GSSD, which is abstract centric, in many ways a document management system seems to be

131

better fitted, as opposed to the Development Gateway, where a relational model, might be a better fit.

This is actually what the Development Gateway uses – a back end relational database (ORACLE) to store its content, while the GSSD uses a hierarchical document management system – the Lotus Domino Workbench.

Content/Editorial Policy:

GSSD has a well-defined workflow for routing of abstracts to content reviewers who approve all content submitted to GSSD. The workflow for an abstract submission within GSSD is reproduced below. Content reviewers review respective language submissions and classify websites using the GSSD ontology. Content reviewers are pre-specified, and are not subject matter experts in any one domain. They are typically, faculty, student research assistants or temporary staff members.

Two features of the GSSD editorial policy bear mention in this context. The GSSD exercises a strict control of abstract submissions to the system as it was originally conceived as a knowledge-based system, not simply an information access device. The GSSD does not allow submissions into its website that are:

- Daily news items, popular magazines or papers, etc.

- Statements of opinion or papers of opinion per se

- Unidentifiable institutional or individual source

- E-materials reporting on other materials whose sources are not known

Another unique feature of the GSSD is a system that is geared to replicate all content in all languages in all mirror sites. This high redundancy strategy (the idea being to let users access the mirror site closest to them, in any language) imposes a constraint on the architecture, and the system has to be geared for high redundancy without inconsistency, and also be capable of rapidly and easily replicating and synchronizing content.

The Development Gateway on the other hand is governed by an Editorial Committee composed of practitioners and experts from government, civil society, and the private sector, who provide guidance on editorial policy, content management models, taxonomy of topics and sub-topics and other policy matters.

Topic pages on the site are managed by guides (individuals and institutions) with demonstrated experience and expertise in the topic area, and are supported by advisors that help evaluate the page, make contributions. Content suggested by users, like GSSD, is reviewed prior to acceptance on the site. Like GSSD multiple languages are supported, however the content is not always available in every language. This eases the burden on synchronization of data amongst the different partner websites, who can now function as more or less autonomous content creation entities.

6.1.3. Nature of Partnerships:

GSSD mirror sites have an arrangement where the partnering institution maintain a server live to the web and integrated into the GSSD system that contains a replica of the entire GSSD database. Although this affords for easy and fast creation of local content, there is very little that can be done in terms of localization of the delivery of content. This is a drawback of the architecture of the system, where once the core application is developed, the regional deployment consists of just translating the 'terms' used in the application to each of the languages. This is managed at present using the Lotus Domino Global Workbench [tm.]

The Development Gateway's partners – in this context individual Topic Page heads, or country gateway portals have considerably more autonomy in creating the format of their own content. The Country Gateways show variations in the knowledge strategy- some are based on the original Gateway structure whereas others appear as 'traditional' portals or directories. At present the content on individual country gateways has not been integrated into the classification system of the main website.

The relative merits of a centralized architecture of GSSD, over a relatively decentralized architecture like Development Gateway's is that it enables better control and is thus ideally suited to an organizational model where individual mirror sites share the task of translating content from all other languages into their own.

However, this architecture does not enable a mirror site to develop its own customized welcome page and additional functionality without impacting the design of the system as whole. Also if strict access control measures are not employed, it is not too difficult for any particular mirror site to tamper with the data of other mirror sites. The policy of translating all content into all languages

also raises the issue of the relevance of content created by any particular mirror site to all other mirrors.

A decentralized architecture with individual Topic Pages or Country Gateways creating their own content and access formats working under a co-branding arrangement with Development Gateway poses its own share of problems. In the context of Country Gateways, while content is mostly local, the Country Gateways and the Development Gateway portal tend to (but not necessarily) have a common information-sharing architecture and editorial policies to ensure that information can be shared globally. In some cases, there has to be a mapping between the content created at each country gateway and the way this content is organized in development gateway.

6.1.4.   Technology Choices:

GSSD's partnership with Lotus, provides it with very specific advantages of:

- Using the excellent collaboration, workflow and document management features with in Lotus Notes;

- The powerful synchronization and replication of content possible through the Lotus Domino Platform;

- Lotus Notes is a RAD (Rapid Application Development Tool), and it is easy to develop and deploy multilingual applications without redesign for each language.

- Is symbiotic, as Lotus is learning a lot about multilingual application development from GSSD, and has dedicated resources, software and support services for use by GSSD.

However, there are certain very serious drawbacks to this software:

a) Proprietary

Content can only be stored as Lotus Notes documents, which is then rendered as HTML through the Domino server. Exporting documents from Lotus Notes to other formats-MSOffice, PDF, XML is not easy, and this might be significant factor to consider when GSSD plans to integrate[why integrate as opposed to just link?] the operational model is one of decentralizastion and distributedness – not integration its knowledge base with other partners.

b) Not Relational

Lotus Notes is a Document management system, which at best can be described as an object database. The predominant design component is a 'document', which could be abstracts, static content, etc. It is difficult to perform even the most basic querying with the

134

system requires creating 'views' or 'agents' –design changes in the system. There are very few third party reporting tools for adhoc querying, and reporting.

c) Performance Issues

The Domino server is not built for high traffic access. Given the loads under which GSSD operates (120 sessions, 10 concurrent sessions –max), the server has proved adequate, but the Domino server has known vulnerabilities at higher loads.

d) Lack of development experience

Availability of skilled technical resources is a key requirement to manage any technological system. It is difficult to find developers and administrators with Lotus/Domino experience. For GSSD, it is hard for the initiative to find skilled Lotus Notes programmers amongst the student community to manage the application.

However, to move away from a proprietary, hieararchical system, one needs an open source multi-site relational database/content management system alternative that will still provide the benefits of easy content management, and multi lingual mirror site replication and synchronization. Although multi-lingual and multi-site clustering solutions for relational databases exist currently- Oracle and MS SQL both provide solutions, they are not linked to content management systems, and there are none in the open source domain. The alternative is to build a system from scratch –using an OS Relational Database- PostgreSQL, and build the content submission, translation and publishing process and repl/synch languages using OS scripting tools.

The Development Gateway's Country Gateway teams in Development Gateway are free to develop and use their own technology platforms. The Gateway system is non-proprietary, and is based on open-source software and standard technology. They offer a platform on which country teams can manage their own content. Hence Country Gateway Teams are encouraged to use the existing platform even if they prefer to customize it for their own special needs. Specifically, the Development Gateway uses, XML for information and data exchange between the development gateway and country gateways, partner organizations, and other systems Tcl, Unicode, UTF8 and plans to adopt: open standard SCORM for e-learning courses, J2EE, SOAP, WSDL, UDDI, WAP and other protocols such as VoiceXML

The Key Benefits to using an Open Source Platform are that it :
- Decreases the cost of sharing and finding information.

- Increases the audience that can be reached.

- Ensures the accessibility and knowledge sharing on a worldwide basis and across users.

135

The key drawbacks in using a completely open source, open standards approach, free for all development being lack of control over the individual development units, the lack of centralized support for all their applications, and the extent of application integration development that might be involved.

6.1.5.   Organizational Structure:

One of the interesting points of comparison to note is the interplay of organizational structure, strategic choices and the technology used. GSSD (with its mirror sites) is an example of a tightly coupled organization, with highly standardized processes that come at the price of high coordination costs and diminished innovation at the mirror sites.

In contrast, there is a high degree of innovation (however, there might be an issue of quality trade offs) at each of the country gateway nodes with local mutations to reflect localized needs. The coordination costs in managing a decentralized network is also lesser.

Another point of interest is the degree of formalization of the relationship. The GSSD relationships with Mirror sites on content are relatively formalized with clearly specified obligations from the partners. In case of the Development Gateway, barring certain exceptions, Country Gateways are flexible in determining the level of involvement they wish to undertake.

However both cases share the similarity of having focused top management involvement that play an active role in determining the day-to-day strategy and in planning the future road map.

It must be mentioned that the comparisons made need to be understood in the context in which the choices were made, and an effort made to appreciate the difference between the kinds of stakeholders involved, the nature of the partnerships, financial considerations, the decision making processes, and the kinds of choices available at the *time* the strategic decisions were taken in the two organizations.

6.1.6.   Potential for Synergy between the Development Gateway and the GSSD:

We examined two different models these two organizations have evolved into with the overall objectives made strategic and technological choices that were very different. There is a lot both can learn from each other. However, the fundamental difference between the two is that while Development Gateway develops and deploys web portal technology and applications as the underlying platform, GSSD operate less with web portal vision than with an integrated knowledge-networking strategy.

Of the four core Development Gateway Foundation (DGF) goals of Increased knowledge sharing, enabling aid effectiveness, building local capacity to empower communities and improving public sector transparency; the first three are very similar to the objectives of GSSD and its parent- the Technology and Development Program (TDP) and this suggests that there are opportunities for mutually beneficial collaborative development partnerships.

In fiscal year 2004, the Gateway's focus for its Topic Pages involves "(1) the decentralization of content management to ensure the value and dependability of content...(2) agree on data structure with partners to create an automated compendium- reinforced by manual quality assurance, and...(3) enhanced knowledge exchange with improved collaborative tools" [95]

MIT's strategy encompasses knowledge sharing, management and development of new knowledge through innovative distributed e-Knowledge management systems, to:

(a)Reduce barriers to access of quality-controlled knowledge in areas critical to development – through the GSSD operating knowledge network of networks;

(b)Build autonomous mirror sites in developing countries for local knowledge provision into global networks and vice versa. Example – Government of China, ACCA21 Office, hosting China-GSSD and supplying the China language system – a total clone of the MIT system;

c)Decentralize content provision through distributed systems of knowledge management, while retaining quality & reliability measures;

d)Provide e- tools for major non-Western languages, such as Chinese and Arabic.

The synergy between the goals of these two different knowledge networks suggest that there are possibilities for collaboration between them. For example through a content sharing partnership between GSSD Knowledge base and Development Gateway's topic pages by introducing selective new content, such as Best Practices and e-Entrepreneurship and strengthening existing content, such as e-Governance and ICT for Development.

Another potential area for collaboration is between GSSD and the Country Gateway network run by the Development Gateway. The country gateway program is in a mature stage where the number of new country gateways introduced have stabilized, and the focus of the program is one of creating an information architecture that specifies the content, layout and connectivity

---

[95] The Development Gateway Business Plan FY03-04: From the Development Gateway website: http://www.developmentgateway.org

standards between individual country gateways and the Development Gateway Portal. The aim of this exercise would be to increase the coordination between and amongst the country gateways themselves and with the main portal.

This initiative would require creating a new knowledge network design that should lay out the information, technological and organizational architecture for increased coordination. This exercise would most likely involve 1. a content audit to develop relevant information products 2. a new design template for the country gateways that specifies procedures for content provision and sharing and layout and 3. a new interface for connectivity at the country gateways including content schema, metadata format, and process workflow for publishing.

Given GSSD's experience in creating and managing a highly structured, yet decentralized network of mirror sites, it could use its existing technology to provide is beyond the scope of this thesis, a brief explanation of what such a system would the Development Gateway with an interface that would enable quality controlled, asynchronous communication (not based on real time data transfer) between the main DG portal and individual country gateways. Although the detailed description of such a 'bridge' or interface system entail is given below.

The proposed model for this interface would be one that enables two way, manually validated, asynchronous communication between any two installations. The purpose of this bridge would be to 'pull' knowledge resources from each of these sites and 'push' it to sites that need it. For example, if one were to envision a bridge between GSSD and the current DG portal, on the GSSD side, the interface would take its feed from the submit site process. The interface would have built in decision rules on where to send this content – like say any content submitted within 'Energy and Environment' within GSSD would go to say 'Environment Law' within the DG portal. A mapping exercise between the current 'slices' within GSSD and the 'Topics' within the Gateway show that there is considerable overlap. The interface would thus able to sort out the content flowing from one partner and intelligently route it to the relevant place in the other partner's ontology. The knowledge throughput of both systems is bound to increase, though it is suspected that the less discriminating partner (in terms of quality of content) is bound to benefit more.

However integration is not an easy task due to the different expectations of participating partners, the difference between organizational processes, the characteristics of members of each network, the technological challenges inherent in integrating proprietary technologies etc. The next section integrates the above learning and attempts to provide an integrated framework to evaluate the performance of any knowledge network. The above brief description for an interface is just one of the many possible models of collaboration between GSSD and the Gateway and

could be generalized to apply to similar collaborations between any two knowledge networks that share the same objectives and recognize the synergy possible in such cross network cooperation.

The next section summarizes our learning so far by integrating the theory behind knowledge networks with our understanding of the ways knowledge networks in practice actually operate. It attempts to develop an integrated framework for evaluating knowledge networks and concludes with some guidelines and recommendations towards steering a knowledge network to sustainability.

6.2. An integrated framework for re-engineering Knowledge Networks:

This thesis adopts an engineering systems methodology in its approach towards characterizing, measuring and evaluating knowledge networks. Engineering Systems are diverse, complex systems that include components from several engineering disciplines, as well as economics, sociology, psychology, and other sciences and the Engineering Systems methodology adopts an inter-disciplinary approach to design, develop, implement, and sustain complex engineering systems. Knowledge Networks are perfect examples of engineering systems, and re-engineering knowledge networks involves a proactive, systematic (and systems based) approach towards planning and design, management, measurement and evaluation to cultivate the conditions for sustained sharing and dissemination of knowledge resources.

As described earlier in section 2.3, knowledge network components are the actors- this includes individuals, groups, organizations; the relationships between actors, which can be categorized by form, content and intensity; the resources which may be used by actors within their relationships, and the institutional properties, including structural and cultural dimensions such as control mechanisms, standard operating-procedures, norms and rules, communication patterns, etc.

Each of these components and the relationships amongst them contribute towards the sustainability of a knowledge network. For instance a successful knowledge network requires both a supportive social atmosphere and an appropriate technical infrastructure to provide it with the same. The technical infrastructure- including IT tool selection, integration, design and training must in term mesh with the control and communications mechanism – like editorial policies, content storage, controlled access and allowed communication channels.

This section attempts to create an integrated framework that helps a network designer or planner layout the objectives and goals of the network, identify the individual components within the

network to describe the relationships among them and establishes interdependencies, and design the technical and organizational architecture using these components and relationships to create a successful knowledge network. The proposed multistaged network thus involves the following phases: Network Planning Phase, the Network Design Phase and Implementation Phase, and the Network Evaluation and Monitoring Phase. A description of the different phases and the framework follows.
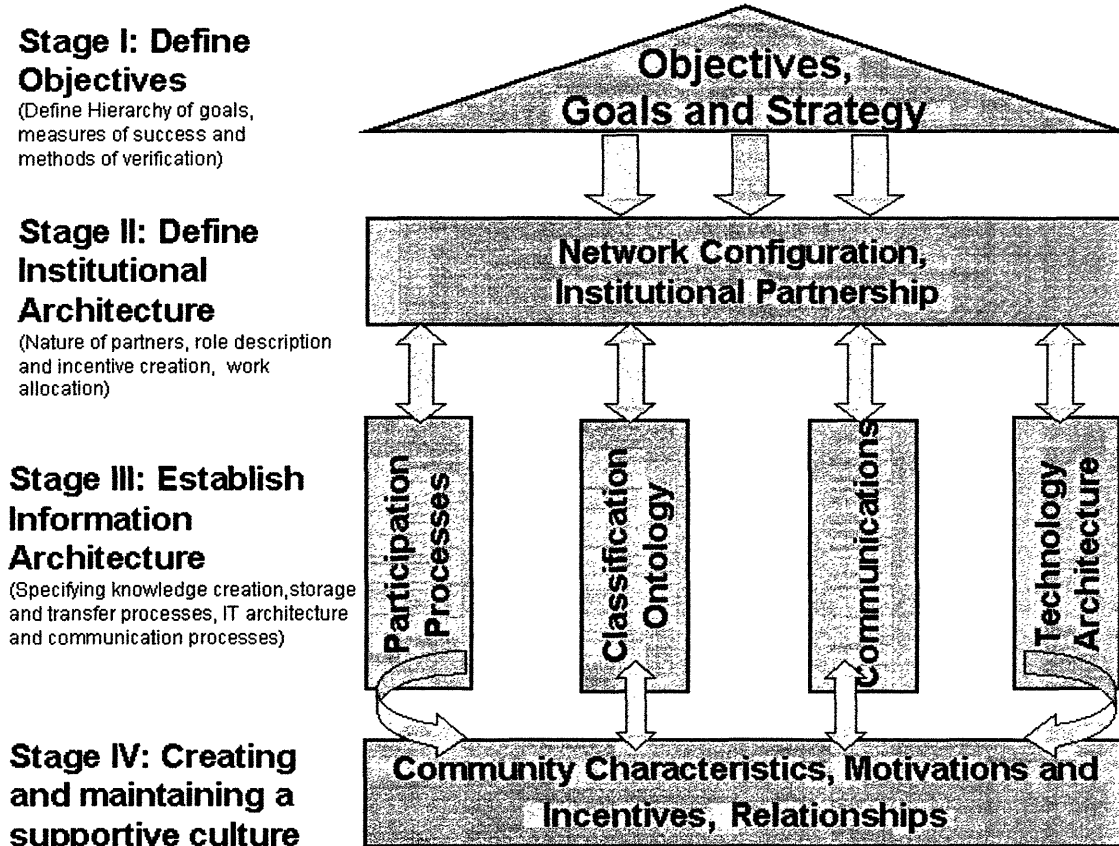
Stage I: Define Objectives
(Define Hierarchy of goals, measures of success and methods of verification)

Stage II: Define Institutional Architecture
(Nature of partners, role description and incentive creation, work allocation)

Stage III: Establish Information Architecture
(Specifying knowledge creation,storage and transfer processes, IT architecture and communication processes)

Stage IV: Creating and maintaining a supportive culture

Objectives, Goals and Strategy

Network Configuration, Institutional Partnership

Participation Processes

Classification Ontology

Communications

Technology Architecture

Community Characteristics, Motivations and Incentives, Relationships

Figure 33: An integrated framework for Knowledge Network Assesment

Knowledge Network Planning Phase: The first stage in the integrated framework involves establishing the goals and objectives of the network. As described in section 2.4, this involves identifying and prioritizing amongst various goals and creating a hierarchy of objectives that takes into account all stakeholder interests –including the needs and objectives of the sponsoring organization, the various institutional partners and the other individuals and groups that participate in the knowledge network. This stage also involves defining operational variables and the processes that reflect the network's ability to reach these goals. Once the hierarchy of goals

140

are defined, the next phase is to choose an organizational and technological architecture most effective in realizing these goals.

Knowledge Network Design and Implementation Phase: This phase incorporates stages II, III and IV of the integrated assessment framework.

Stage II involves laying out the organizational architecture that decides the number and type of partners, the level of control by the sponsoring organization, the degree of centralization, task assignment and alignment of incentives of partners, the amount of workload per partner and the network governance mechanism. Stage III is concerned with the planning of the information architecture and lays out the set of principles and standards that guide the high level design, selection, construction, implementation, support, and management - of the organization's processes and communications infrastructure. Stage IV involves defining and maintaining an network culture (including amongst others the values, purposes, structural relationships, language, etiquette and history of the network) - and engineering trust, tolerance and rewards systems in place to reinforce the culture.

This phase involves both planning and the actual implementation and often includes creating low-level designs, work schedules, performance targets and relevant incentives, rules of participation, moderation and control.

Knowledge Network Monitoring and Evaluation Phase: The monitoring and evaluation phase is an ongoing activity that involves setting up both the implementation of performance measurement processes and the decision rules to act on them. This phase specifies who has to monitor what and when, and what to do with the results. The organization develops and trains its personnel in the different methods of analysis (traffic analysis, content analysis and network mapping, surveys etc.) and provides guidelines on how to interpret and act on the results.

It has to be mentioned here that knowledge networks as living systems also exhibit features of adaptability and emergent behavior and do not always follow a plan-design-implement cycle. Very often processes, norms and relationships are created within a network that were not originally envisoned. The framework helps in deciding on which kind of emergent behavior needs to be encouraged and adopted and what needs to be censored or dropped.

The integrated structure outlined above uses an engineering systems approach to provide both a micro and a macro perspective that enables us to comprehend knowledge networks in its entirety. The multi phase framework enables laying out the individual components while recognizing the interdependencies that exist among them. For example, the framework recognizes that networks

are structural as well as cultural, and that the planning, design and the evaluation phase will have to take into account the individual-organizational, the technological-institutional and the individual-technological inter-connections that exist.

6.3. Summary and Conclusions:

6.3.1. Summing Up:

The primary aim of this thesis is to help knowledge network researchers and managers understand knowledge networks from an integrated perspective, identify barriers to sustainability and to recommend re-engineering options.

The first step towards understanding knowledge networks is to classify them. This thesis provides a classification framework to categorize knowledge networks according to their varying objectives, facilitating conditions, work processes, network architectures, membership structures, and institutional configurations. The importance of knowledge networks within the domain of sustainable development and the unique characteristics of such networks were also explored.

The thesis then defines the components of 'success' and 'sustainability' within a knowledge network and looks at the different objectives underlying different networks. We saw that clearly specifying objectives and goals, with the operational variables to measure them is key. Knowledge networks usually have multiple goals, and we saw the importance of creating a hierarchy of goals according to their importance. To measure the performance of a network in achieving this goal, we constructed a two-stage evaluation framework where the first stage attempts to identify the effectiveness of a network in reaching its objectives, and the second stage looks at the efficiency of the network in reaching these objectives.

Our characterization of knowledge networks as inter-linked subsystems led us to explore the human, technological and institutional subsystems where we explored the motivations and incentives for collaboration within a knowledge network. We studied the existing literature on motivation, compared case studies of motivations and incentives in three resilient, successful configurations of networks - corporate knowledge management systems, open source projects and P2P file sharing systems. We further explored the problem of under provision and absence of collective action, due to free riding. We found that solving free riding requires solutions that involve devising selective incentives, changing network characteristics to achieve critical mass, or making changes to the institutional and organizational structure.

We then looked at the application of these theories on two knowledge networks in the real world-
the GSSD and the Development Gateway and compared and contrasted their goals, approaches,
processes and network architectures. We identified rich opportunities for collaborative
development between these two knowledge networks and provided recommendations on
interfaces between these two systems.

The theories, frameworks, case studies and analysis approaches used were then synthesized to
develop an integrated framework for assessing knowledge networks to serve as an aid to a
knowledge network manager in the planning, design, implementation and evaluation phases.

6.3.2.   Next Generation Knowledge Networks, Areas for further research, and Conclusions:

Knowledge Networks are complex beasts. The large number of component systems, the inter-
linkages amongst them, the uncertainty in their emergent behavior from the interaction with its
environment, makes it difficult to design and manage. Knowledge Networks thus fall under the
category of Sussman and Dodder's CLIOS [96](Complex, Large, Integrated and Open Systems)

The study of knowledge networks thus requires a new breed of researchers capable of a
sophisticated understanding of the complex socio-technical processes involved. This requires the
adoption of an interdisciplinary systems based approach – one that combines a deep
understanding of each of the disciplines (information and communication technologies, sociology,
organizational psychology and political science) and an appreciation of the nature of the
interrelationships amongst these disciplines.

This thesis uses the Engineering Systems Approach towards understanding knowledge networks
by using an integrated analysis of the interactions of the technologies, individuals and institutions
from a systems perspective. By looking at the various subsystems involved - human and
organizational (the individual, network and institutional level), the technological subsystem and
the social subsystem this thesis develops a framework for looking at the knowledge network as
an integrated whole.

Although the approach adopted this thesis could be used for analyzing most configurations of
knowledge networks and application domains, this thesis is particularly interested in the
characteristics of knowledge networks in the sustainable development domain. The interest in
building and managing successful knowledge networks in development is particularly high among

---

[96] Sussman, Joseph & Dodder, Rebecca, 2002. 'The Concept of a "CLIOS" Analysis Illustrated By
The Mexico City Case.' Prepared for MIT's ESD Internal Symposium.

International Development Organizations (IDOs) who view these networks as an important part of their capacity development activities. Knowledge sharing at these IDOs has evolved over time- from an emphasis on capturing and organizing knowledge, to their current focus on adopting, adapting, and applying knowledge and on connecting knowledge workers.

However, knowledge networks are continuously evolving and the next generation of knowledge networks are likely to be those that go beyond the sharing of knowledge to those that are able to translate knowledge into action. In the context of sustainable development, this would involve integrating knowledge principles and practices into the institutions and policy frameworks that are shaping the information society. Evidence of this phenomenon is already around us- from flourishing best practices networks, to the shift towards Just In Time Knowledge- where knowledge-sharing practices are integrated into organizational work processes. The next generation knowledge network would also be a more 'intelligent system', through the development of dynamic and adaptive ontologies for representing continuously changing knowledge domains, using data mining tools for analyzing linkages for improve cross-disciplinary understanding etc.

Technological improvements in the form of the next generation internet, the semantic web and the growth of web services are likely to change the notion of knowledge storage and dissemination as we know of it today, allowing us to transform knowledge by combining, classifying, and analyzing it in new ways and with the emphasis shifting from connectivity to new levels of interactivity, would allow us to learn and create across disciplines, languages, & cultures.

It is hoped that the methods of analysis described, the case studies discussed and the integrated framework developed would aid knowledge network managers in charge of re-engineering knowledge networks perform better. It is expected that these frameworks, models and approaches will have to be refined as the nature of knowledge networks themselves continue to change.