Evaluation Findings of PAN Localization Project

Sana Shams
Mudasir Mustafa
Atif Mirza
Yasmeen Daud
Qaisar Khalid Mahmood
Sarmad Hussain

Center for Language Engineering (CLE)
Al-Khawarizmi Institute of Computer Science (KICS)
University of Engineering & Technology (UET)

www.cle.org.pk

www.idrc.ca

Acknowledgements

Acronyms

ACSA	Afghan Computer Science Association
ASR	Automatic Speech Recognition System
ASIWG	Arabic Script IDNs Working Group
BWN	Bangla WordNet
BPPT	Agency for Assessment and Application of Technology
BIPS	Bhutan ICT Policy and Strategy
BSG	Broadband Stakeholders Group
CAN	Computer Association of Nepal
ccTLD	Country Code Top-Level Domain
CRULP	Center for Research in Urdu Language Processing
CRBLP	Center for Research on Bangla Language Processing
CLE	Center for Language Engineering
CMS	Content Management System
CPIs	Country Partner Institutes
CRC	Communications Regulatory Commission
D.Net	Development Research Center
DIT	Department of Information Technology
DVD	Digital Video Disk
ENRD	E-Network Research and Development
EUL	End-User Level
FOSS	Free Open Source Software
gTLD	Generic Top-Level Domain
OMg	Gendered Outcome Mapping
GOL	Government of Laos
GMT	Gender Mainstreaming Tool
GPI	Gender Parity Index
GSK	Language Resource Association
HLT	Human Language Technology
HLCIT	High Level Commission on IT
HMM	Hidden Markov Model
HTK	Hidden Markov Model Toolkit
HTML	Hyper Text Markup Language
ICT	Information and Communications Technology
ICANN	Internet Corporation for Assigned Names and Numbers
IDNs	Internationalized Domain Names
IDRC	International Development and Research Center
IT	Information Technology
ITC	Institute of Technology Cambodia
IDNA	Internationalized Domain Names in Applications
ICTA	Information and Communication Technology Agency
ICTPA	Information and Communications Technology and Post Authority, Mongolia
LLC	Local Language Computing

LLT	Language Learning Tool
LIP	Language Interface Packs
LLL	Local Level Leader
LLLT	Local Level Leader Training
LTK	Language Technology Kendra
LLWG	Local Language Working Group
MoIC	Ministry of Information and Communication, Bhutan
MCIT	Ministry of Communication and Information Technology, Afghanistan
MoEYS	Ministry of Education, Youth and Sports, Cambodia
MT	Machine Translation
MUST	Mongolia University of Science and Technology
MoITT	Ministry of Information Technology and Telecommunications
M&E	Monitoring and Evaluation
MPP	Madan Puraskar Pustakalaya
MoU	Memorandum of Understanding
MLP	Microsoft Localization Program
NECTEC	National Electronics and Computer Technology Center
NiDA	National ICT Development Authority of Cambodia
NITP	National Information Technology Park
NAST	National Authority for Science and Technology
NUM	National University of Mongolia
NUCES	National University of Computer and Emerging Sciences
NLP	Natural Language Processing
NLIT	Nepali Language in IT
NCGA	Nepali Computational Grammar Analyzer
OSS	Open Source Software
OCR	Optical Character Recognition
OS	Open Source
PTA	Pakistan Telecommunication Authority
POS	Part of Speech
RCB	Research Capacity Building
R&D	Research and Development
RIB	Research Initiatives Bangladesh
SMT	Statistical Machine Translation
SLP	Statistical Language Processing
SPSS	Statistical Package for Social Sciences
STEA	Science Technology and Environment Agency
TnT	Trigrams'n'Tags
TTS	Text to Speech
TOT	Training of Trainer
UCSC	University of Colombo School of Computing
UI	University of Indonesia
UPS	Uninterruptible Power Supply
XML	Extensible Markup Language

Table of Contents

1 Introduction

The use of information communication technologies (ICTs) is often accentuates as the most promising and fundamental driver of social and economic development. Target 18 of the Millennium Development Goals prescribed by the United Nations (2006) stipulate that "*In cooperation with the private sector, make available the benefits of new technologies, especially information and communications.*"

Aiming to achieve this impact, the developing countries are rapidly deploying ICTs especially the Internet to make improvement in social and economic conditions and enhancing the standards of people's life (Batchelor- et al, 2003).

ICT Facts and Figures compiled by Internet Telecommunication Unit in 2011 reveals that while 35% of the total world's population is using the internet, internet access in Asia is about 26% as compared to 78% in North Americas and 61% in Europe. This low ICT penetration in Asia is not only because of the unavailability of infrastructure, but also because computing is still primarily in non-Asian languages (Hussain and Mohan, 2008).

PAN Localization Project was conceived as a regional initiative to develop local language computing capacity in Asia specifically focusing to develop indigenous human resource capacity for undertaking localization research across both genders, develop localized technology for Asian languages, and advance policy for local language content creation, access and use across Asia. This report presents the evaluation results of this project and sets forth relevant recommendations based on the data analysis. The evaluation study presented is both retrospective and prospective. It evaluates the project outcomes in perspective of the project's goals established at its inception and also discuss the prospects of how PAN localization project has contributed to building the national capacities their effectiveness and future sustainability of its capacity development support.

The report is organized in seven chapters. Chapter 1 introduces the report while Chapter 2 presents the scope and purpose, lists the objectives of the project and summarizes literature related to the project's evaluation objectives and domains. Chapter 3 presents findings on research capacity building in participating countries through PAN Localization project. Chapter 4 presents the data on preference for and impact of language for learning the use of ICTs in participating countries. Chapter 5 discusses the PAN Localization project influence on local language computing policy at regional and country levels. Chapter 6 discusses efforts of project partners in mainstreaming gender in PAN Localization project. Chapter 7 sums up the experiences of PAN Localization and present recommendations based on the lessons learnt for further implementation, in the context of localization.

2 Scope and Purpose

Monitoring and evaluation (M&E) of development activities provides government officials, development managers, and civil society with better means for learning from past experience, improving service delivery, planning and allocating resources, and demonstrating results as part of accountability to key stakeholders. World Bank, 2004[1]

2.1 Background

Phase I of PAN Localization project (2004-2007) focused on developing local language standards and technology across seven Asian countries. The countries (and languages) included in the project were Afghanistan (Pashto), Bangladesh (Bangla), Bhutan (Dzongkha), Cambodia (Khmer), Laos (Lao), Nepal (Nepali) and Sri Lanka (Sinhala, Tamil). During this phase the project undertook major work in the area of HR capacity building in local language computing by conducting extensive seminars and workshops for its partners and hands on research to develop local language software like localized keyboard maps, local collation standards and fonts; terminology standards and open source operating systems in Nepali (Nepalinux) and Dzongkha (DzongkhLinux); working systems for OCR for Bangla, Lao and Sinhala; lexica, spell checker for Bangla, Dzongkha, Khmer, Nepali, Sinhala and Tamil and text to speech systems for Sinhala.

Phase II of PAN Localization project (2007-2010) was extended to ten countries with a more specific focus to plug the localized technology developed in project's Phase I into usable applications from which end-user groups like rural students, teachers, print-disabled community, monks, government officials, rural housewives, etc. This includes development of localized technology interfaces and preparing localized training resources (both software and published material) for training the community on how to effectively and efficiently use the localized ICTs to serve their needs. The project also focused on local language content development by developing models to trigger content generation and consumption by the end-users. End-Users groups in target countries are being trained on developing online content, while where necessary; collaborations are sought with resource partners, health ministry's for health information, lawyers for legal document, and other government sectors for bringing information online. At the regional level, research units focused on compilation and synthesis of effective end-user training methodologies, content related IPR issues and resulting policy implications and localized computing frameworks to help other nations to develop their language specific indigenous solutions.

The purpose of conducting the evaluation is to provide an objective, disciplined and systematic assessment of the second phase of PAN Localization project. The evaluation is intended principally for learning purposes. It is expected to generate relevant findings, lessons, and recommendations which will be shared with key stakeholders of the project and used by the donor agency to guide and inform future programming. This evaluation is formative in nature as it aims to seek improvements in the current projects design and accomplishments.

---

[1] http://www.worldbank.org/oed/ecd/

2.2 Intended Users and Uses

Following the utilization focused evaluation's philosophy, this evaluation study focuses on drawing successful lessons from the evaluation results to gather both strengths and weaknesses of the program to foster reflection and innovation in future project designs. In this context the intended users of the evaluation are the project donors, IDRC, the project staff specifically the regional project leader based at the project secretariat at CLE and country project team leaders and the gender evaluation methods project team.

Specific intended uses of evaluation for each of the intended user are described below:

Project Donor:

- To understand how effective PAN Localization program has been in achieving its project goals and objectives
- To assess the relative impact of PAN Localization project in developing sustainable capacity for R&D in local language computing
- To study if PAN Localization project has been able to impact policy in any of the partner countries

Project Staff:

- To study how effective has PAN Localization project been in reducing the digital divide through development of localized solutions

Gender Evaluation Method's Team:

- To study how effectively gender has been mainstreamed throughout the project life cycle

2.3 Evaluation Objectives

The evaluation study follows the principles of a utilization focused project evaluation. The study is a formative assessment of the project implementation and strategies practiced across the country project teams, premeditated to better inform future programming decisions.

Broadly the project evaluation aims to:

- gauge the project contribution to raise the technological support, and to nurture a sustainable capacity for R&D in local language computing for the target Asian languages
- study the impact of the project in advancing digital literacy in the target communities
- study how far has the project contributed in influencing the local political machinery to be more receptive for conducting local language computing projects
- assess how far has gender concerns been mainstreamed throughout the project life cycle

Based on the evaluation objectives, the following evaluation questions and sub questions have been established.

2.4 Evaluation Questions

1. To what extent has localization played a role in promoting digital literacy in the target end-user communities?

   a) To what extent has localization enabled the target end-users' accessibility to ICTs?
   b) To what extent have localized ICTs enabled communities to retrieve online content?
   c) To what extent have localized ICTs enabled communities to generate online content?
   d) How effective was localized technology in enabling end-users to generate online content?

2. To what extent has PAN Localization project contributed to develop sustainable capacity for R&D in local language computing?

   a) To what extent has PAN Localization project contributed to improve individual's capacity for R&D in local language computing?
   b) To what extent has PAN Localization project, focused to improve organization's capacity for R&D in local language computing?
   c) How far has the project contributed to develop/strengthen research centers for R&D in LLC?

3. To what extent has PAN Localization project contributed to influence local language computing policy?

   a) To what extent has the project affected public policies, laws, regulations, programs and government structures?
   b) How far has PAN Localization project contributed to influence the policy development process by improving the intellectual framework?
   c) How far has the project contributed to the development of the network of technologists, social scientist and policy makers for collaborative work on policy development?
   d) To what extent has PAN Localization project, influenced policy development to raise sustainable capacity for R&D in LLC?

4. How far has PAN Localization project been able to ensure equitable participation of both men and women in the project?

   a) What special needs of both men and women for localized ICT's have been considered in the project planning and design?
   b) To what extent has the project taken measures to ensure equitable participation of women in the project team?
   c) How far has the project been successful in achieving gender positive outcomes?

2.5 Research Capacity Building in Local Language Computing

Local language computing also called localization or enabling ICTs in a local language is essential for effective ICT use in Asia (Gul, 2004). As defined by Hussain and Mohan (2007), localization is "The *process of developing, tailoring and/or enhancing the capability of hardware and software to input process and output information in the language, norms and metaphors used by the community*." It is a three step process. First, the linguistic analysis is required to document (and standardize) language conventions that are to be modeled. Second, localized applications (both basic and intermediate level) e.g. fonts, keyboard, locale, spell checkers, etc. need to be developed to enable basic input and output of text in a local language. Thirdly, to provide comprehensive access and assist content development, advanced applications like translation systems, speech dialogue applications, etc., need to be developed. Localization therefore requires significant knowledge of linguistics (phonetics, phonology, morphology, syntax, semantics and pragmatics), signal and speech processing, image processing, statistics, computational linguistics and advanced computing (Hussain et al, 2007). This research being language dependent, entails nurturing indigenous *research capacity* (Breen et al, 2004, DFID, 2010) at the levels of individuals, organizations, and systems to sustain.

Capacity Building within the context of research is enhancing the abilities of individuals, organizations and systems to undertake and disseminate high quality research efficiently and effectively (DFID, 2010). "*Capacity building is a process whereby people are enabled to better perform defined functions either as individuals, through improved technical skills and or professional understanding, or as groups aligning their activities to achieve common purpose*" (Breen. et.al, 2004).

Strengthening indigenous research capacity versus technology transfer (Harris, 2004; Nokolov and Illieva, 2008) is the most effective process for advancing research, specifically in localization. However building scientific research capacity in developing countries has a history of posing challenges at various levels including HR, research priorities and funding, research facilities, communication, import policies for infrastructure (Vose and Cervellini, 1981). Even today, conducting research in developing countries pose challenges akin to those that had existed before. These include material constraints including scarcity of financial resources, lack of appropriate technical infrastructure for conducting research as well as a less thriving trend for advancing scientific careers due to lack of scientific research tradition, institutional support and collaboration within the local scientific community. This is also perpetuated due to the fact that available academic training, market demand and placement is often development oriented in nature rather than research oriented (Harris, 2004). Manuscript writing is another research area which is often neglected. Scientific publications in peer reviewed journals significantly improve the credibility of the research conducted and also help promotes its visibility and dissemination value including improving chances of obtaining funding.

Capacity building thus needs to take place at individual level as well as at the institutional and policy environmental level at large. Also, in order to address capacity building challenge holistically, interventions must not only target skill development in the technical discourse but must also include other equally significant aspects including development of research infrastructure, linkages and collaboration and focus on conducting close to practice research and dissemination and impact monitoring.

2.6 Localization Research and its impact on Policy

Knowledge utilization is the prime focus of a research project through either its application in technology or in decision making and policy. The knowledge generated thenceforth is useful in improving societal conditions (Carden, 2004). Development research projects aim to build the capacity of the indigenous people in effective utilization of the developed technologies and techniques and also specifically focus to address the contemporary policy. Some research projects are conducted with the aim of very immediate impact and focus their research on issues of direct relevance in their societies. While some research projects are launched to those issues that are very significant but hidden. The impact of such type of research project may not be as visible as the former one, but these projects contribute more in the development of the country. Policy makers often utilize research outcomes in decision-making however different perspectives on research utilization in policy development exist that are explained below.

Caplan (1979) developed the theory of "*two communities*" that discuss the use and non-use of research in policy development. They elaborate research utilization as a symptom of the cultural and behavioral trait, a gap between researchers and policy makers. According to him, both researchers and policy makers have different perspectives and it becomes the cause of limited utilization of research in policy development. In addition, the notion of a cultural gap between researchers and policy makers has positively explained the minor utilization of research in developing policy.

While Caplan (1979) believed that research outcomes are directly involved in policymaking process however Weiss (1977), Webber (1991) and Sabatier and Jenkins-Smith (1993), citied by (Neilson, 2001) do not believe this relationship as linear. They think that research-policy link is not a direct one, particularly in relationship to data and information sources. To them, research is among one of many sources of information for policy makers, and that it is not a simple dichotomy between use and non-use but rather that knowledge/research utilization is built on a gradual shift in conceptual thinking over time (Neilson, 2001).

According to Weiss (1977), research has mostly no direct relevance to policy decisions, but can make contribution in influencing policy through other important ways, namely by changing the language and perceptions of policy-makers and their advisors. Such change may occur less decisively, but will exert influence in a powerful manner over a longer period of time. Weiss refers to this perspective as the "enlightenment". Through this approach, research contributes indirectly over time through the circulation and "percolation" of ideas and concepts in influencing policy (Lindquist, 2001).

How research can be used in influencing or making policy is the fundamental question that needs to be addressed when any research is being considered to influence on policy. Weiss (1977) has discussed seven ways through which research can contribute in policy development. These points are elaborated below.

1. Knowledge-driven: application of basic research; this model assumes that basic research provides an opportunity for policy-relevant research which can then be applied;

2. Problem-solving: communication of research on an agreed upon problem to the policy maker; this model implies that there is consensus between the researchers and the policy makers on the solution or end-state;
3. Enlightenment: education of the policy maker; that with time the accumulation of research will influence policy by educating the policy maker;
4. Political: rationalization for previously arrived at decision; used by policy makers to bolster support or provide ammunition for opposition;
5. Tactical: requesting additional information to delay action; often used by government agencies or other organizations/institutions as a response to a problem or issue;
6. Interactive: competing information sources; this implies that policy makers are actively searching for policy-relevant information that is not based on social science research; this type of use is considered to be more realistic of how policy makers use information in the policy process;
7. Intellectual enterprise: policy research is just one type of many intellectual pursuits (Webber, 1991; Auriat, 1998 citied by Neilson, 2001).

According to Lindquist, 2001, "*Assessing policy influence, then, is typically about carefully discerning intermediate influences, such as expanding capacities of chosen actors and broadening horizons of others that comprise a policy network*." This requires developing a full view of the range of actors involved in a project's "domain", the nature of relationships among those actors, and a very good sense of how that network and policy field has evolved over time".

2.7 ICT Literacy and Local Language Computing

The digital divide refers to the gap between those who have access to information and computing technologies (ICTs) and effectively benefit from technology, and those who do not (Arrison, 2002: Ryder, 2006). Many factors have been identified that are deemed to widen this digital divide, including (i) physical access (being able to find a computer to use), (ii) financial access (being able to afford a computer), (iii) language savvy (being able to understand English the lingua franca for ICTs), (iv) education (being able to get benefits from technology), (Keniston 2004, Arrison, 2002).

The digital divide has often been measured in terms of technological access rather than the use of technology and development of relevant content to foster effective use, even though the divides for the latter two are likely to be greater than the access divide. Thus to reduce the digital divide, barriers beyond mere access also need to be addressed to give all users the potential to benefit from ICTs. As reported by the (European Commission, 2009) among the top barriers to ICT uptake are a lack of perceived need; cost (of equipment and/or service); and lack of skills to use ICTs salient issue being the inability to understand English.

Ethnologue (2005) asserts a list close to 7000 languages around the world, off which more than 2600 are spoken in the Asian region, while, English, the lingua franca for ICTs, is understood by less than 15 per cent population of the World (Prado, 2010). Linguistic divide is widely observed on internet as more than 65% of the content on internet is in English while majority (59.8 percent) of the total global online population comes from non-English speaking zones (Sood, 2002). Thus access to information available in

English language is a significant obstacle in learning to use ICTs and empowering people with the rich information available on internet.

Research also clearly indicates that children who study in their mother tongue usually learn better and more quickly, participate actively in the classroom and demonstrate greater self-confidence than children studying in a second language (Cueto et al., 2009; UNESCO, 2008; Benson, 2004; Woldemikael, 2003).The IEA SITES study (Law et al., 2008), which surveys 22 participating educational systems world-wide, reports that "language [is] an obstacle for schools in ICT-implementation in teaching and learning" where English is not the primary language spoken at home, English being the lingua franca for ICTs.

The usage level of the internet and benefits that can be obtained from using the internet depends not only upon access to internet but also on the level of English language skills and availability of local content. Research indicates that level of English influence the use and usage level of internet and non-English speakers like Asia face difficulties to interact with the virtual global community (Lambert, 1996; Horrigan, 2010, ITU, 2010).

Both access to computer and use of internet has increased in the past decade. Yet, according to the statistics presented by ITU (2011), it is observed that the gap between the developed and the developing countries has widened over the same period of time. For instance, in 2002 the proportion of households with computer in Bangladesh was 0.8, which gradually increased 1.9 in 2007 and 3.1 in 2010. In Nepal, the proportion of households with computer was 0.4 in 2002 and 4.2 in 2010. While, in US, which is a highly developed country, the proportion of households with computer was 59, that increased to 70.2 in 2007 and 72.5 in 2010. These statistics show the gap between the highly developed countries and the developing world remains enormous.

Enabling local language computing is essential for access and generation of information, and also urgently required for development of Asian countries. This need was first highlighted in the World Summit on the Information Society (WSIS) meetings held at Geneva 2003 and Tunis 2005, where specific targets were developed focusing on local language software and local language content. They were (i) *n*urture the local capacity for the creation and distribution of software in local languages, (ii) support local content development, translation and adaptation, digital archives, and diverse forms of digital and traditional media by local authorities, (iii) provide content that is relevant to the cultures and languages of individuals in the Information Society, through access to traditional and digital media services and (iv) enhance the capacity of indigenous peoples to develop content in their own languages.

2.8 Gender Mainstreaming in ICT projects

Gender is a social construct that defines roles, behavior, activities and attributes associated with being male and female (United Nations, 1995). These behavioral norms are learned through the process of socialization and are subject to modification with the passage of time. In most of the societies, roles and behavior assigned to men and women differ significantly. Patriarchal systems, e.g. lead men to control over resources, opportunities as well as decision-making (Coomaraswamy, 2005). Women remain undermine leading to gender differences and inequalities. However, in recent years, efforts have been

made to empower women and give equal status, opportunities and access to the resources as available to men by strongly emphasizing gender equality within each of the Millennium Development Goals (MDGs). Thus development projects must ensure gender equality within the program designs and execution (Doran, 1990) principally by addressing gender concerns like, equality of benefits for women and men. Gender mainstreaming approaches thus need to be adopted to ensure gender equality in project outcomes.

According to the United Nations (1997) declaration, "*Mainstreaming a gender perspective is the process of assessing the implications for women and men of any planned action, including legislation, policies or programmes, in all areas and at all levels. It is a strategy for making women's as well as men's concerns and experiences an integral dimension of the design, implementation, monitoring and evaluation of policies and programmes in all political, economic and societal spheres so that women and men benefit equally and inequality is not perpetuated. The ultimate goal is to achieve gender equality.*"

Gender mainstreaming is the process of bringing a gender perspective into the mainstream activities at all levels, including in policies, programmes and projects. It appeared for the first time in international texts after the United Nations Third World Conference on Women (Nairobi, 1985), in relation to the debate within the UN Commission on the Status of Women (CSW) on the role of women in development. It was seen as a means of promoting the role of women in the field of development and of integrating women's values into development work. After Nairobi, international development agencies and governments promoted mainstreaming as a new strategy for taking women's concerns into account. By bringing women's issues into their mainstream policies, programmes and projects, they hoped that earlier problems of marginalization would be overcome.

In addition, United Nations Fourth World Conference on Women held (1995) developed a framework of gender mainstreaming. The major points of that framework discussed below:

1. The initial identification of issues and problems across all area(s) of activity should be such that gender differences and disparities can be diagnosed
2. Adequate accountability mechanisms for monitoring progress need to be established
3. Assumptions that issues or problems are neutral from a gender-equality perspective should never be made
4. Gender analysis should always be carried out
5. Clear political will and allocation of adequate resources for mainstreaming, including additional financial and human resources if necessary, are important for translation of the concept into practice
6. Gender mainstreaming requires that efforts be made to broaden women's equitable participation at all levels of decision-making
7. Mainstreaming does not replace the need for targeted, women-specific policies and programmes, and positive legislation; nor does it do away with the need for gender units or focal points

Whenever women or men are in a particularly disadvantageous position, this perspective addresses gender-specific activities and affirmative action for gender balance. Gender-specific interventions can

target women exclusively, men and women together, or only men, to enable them to participate in and benefit equally from development efforts (ILO Evaluation Unit, 2000)

Gender equality and equity within projects require strategic planning for setting out gender responsive policy, objectives, action, time frame and resources (Murison, 2004). Swedish International Development Agency (SIDA) has identified gender mainstreaming strategies to largely contribute at three levels; in organization's policy framework, its program activities and the project outcomes focusing on its impact upon the community (Shalkwyk et al. 1996).

While specifically focusing on gender mainstreaming in ICTs related projects, gender integration approaches or frameworks are required for systematic implementation. This is equally relevant for ICT localization initiatives aswell. In this context USAID developed a gender integration approach that addresses three major phases of the project namely i) Project Assessment, detailing the policy situation and needs analysis, ii) Project Design, and iii) Project Implementation including HR development and management, technology development, monitoring and evaluation, and service, product development and delivery. Similarly, The World Bank has also provided a checklist of key points to be considered for incorporating gender perspective in ICT related projects. This checklist is based on general principles of gender analysis that are common to all projects as well as specific issues pertaining to ICT project implementation. The first part of this checklist contains questions related to the inclusion of gender issues in the project cycle of ICT projects or projects with significant ICT components. The second part outlines indicative gender issues found in ICT projects and components and includes: national strategies for information infrastructure, strengthening of regulatory structures, and ICTs in education and training.

3 Research Capacity Building in Local Language Computing

PAN Localization project has been focused on research capacity building (RCB) in its collaborating partner countries. The project approached this challenge in a bottom up fashion. The project facilitated development of research units within each partner organization. It then provided support to these research groups by providing funding to meet recurring administrative expenses, technical training through experienced mentors to accomplish research targets, exposure and linkages with regional experts in the field, thus providing the right environment for thriving and institutionalizing local language computing research in respective countries. The following section presents the evaluation framework used for measuring research capacity built in the partner countries through PAN Localization project.

3.1 Framework for Evaluation of Research Capacity Building (RCB) through PAN Localization project

RCB frameworks available in literature (Cooke 2005; Neilson & Lusthaus 2007; Wignaraja 2009) largely recommend three *structural levels* and six basic *principles* upon which research capacity building interventions must focus.

Structural levels define the point of view upon which capacity development initiatives must be targeted. They include *individual*, *organizational* and *system* levels (Neilson & Lusthans 2007, Breen et al, 2004). Some RCB frameworks recommend a hierarchal categorization of the structural levels (Potter & Brough, 2004) according to which capacity building interventions focused at a certain level necessitates capacity development at the prior level. Interventions however cannot be carried out at a certain level in isolation as focus on capacity building at individual levels does have impact on the organization in which it is present. Every activity accomplished at a certain level has impact on the other levels.

Principles of capacity building define the thematic categorization upon which RCB must focus at a certain structural level. Cooke (2005) recommends six principles of capacity building that are listed below:

1. Skill Development
2. Training on Close to Practice Research
3. Development of Linkages
4. Infrastructure
5. Dissemination and Impact
6. Sustainability and Continuity

Thus for the purpose of evaluation, research capacity building (RCB) through PAN Localization project has been evaluated upon these six principles by synthesizing data on project specific indicators defined for each principle as discussed in the section below.

3.1.1   Principle 1: Skill Development

Research capacity building is a multi-faceted process that requires training and supervision to develop technical, managerial, and publishing skills (Harris 2004; Raina 2007). This skill development can also be instrumental in career development and generating opportunities to apply research in practice (Rhee and Riggins. 2007). To evaluate skills development through PAN Localization project the following indicators were used.

1. Completion of project software deliverables
2. Publications of technical papers on project research

Country teams had developed a project plan and agreed to complete a set of project deliverables based on the contract. Successful completion of these deliverables was considered as an indicator of research capacity development. This is because targets to complete project deliverables simulated a real time research scenario enabling researchers to work on real localization problems and find research solutions, through involvement in problem identification, project designing, implementation, quantitative and qualitative analysis thus significantly improving their capacity to conduct research.

Ability to publish research in the form of research papers is a salient indicator for measuring the researcher's research capacity. Thus the number of research publications produced by country teams at various national as well as international research conferences has been used as the second indicator for analyzing research capacity enhancement.

3.1.2   Principle 2: Training on Close to Practice Research

Research capacity building entails the ability to produce research that is useful for informing policy and practice (Cooke, 2005). Research work is "close to practice" when new knowledge generated can directly impact development. Following this principle, the project improved the capacity of the teams to conduct close to practice research by initiating a research focus on deployment of technology within the end-user communities.

Data was thus collected on the following variables from the project beneficiaries in communities of Pakistan, Nepal, Cambodia and Bangladesh where this research work was being conducted, to show the relevance of this research to the practical needs of the communities. These four indicators were:

1. Proficiency of end users in the language spoken at home
2. Proficiency of end users in the language spoken at work
3. Reading skill of end users in English vs. local language
4. Writing skill of end users in English vs. local language

3.1.3   Principle 3: Development of Linkages

Research groups often operate in isolation, limiting the scope and success of their work. Thus in order to enhance the research capacity, resources must be appropriately linked up and connected with active groups working on similar initiatives for robust and collaborative learning. Experiences of researchers

who are working successfully under similarly resource-constrained conditions engender trust and motivation.

Developing linkages, partnerships and collaborations is a reciprocating process of involving organizations in the knowledge information chain for fostering development and diffusion of quality research (Wignaraja, 2009; Breen et al 2004). It also harnesses an increased knowledge base for research development and enhancement. The following indicators have been used to determine the extent to which country project teams have focused on developing linkages, partnerships and collaborations.

1. No. of formal organizational collaborations
2. Participation of teams on research groups

Partner teams were encouraged to establish partnership and collaboration with institution that had more expertise in a specific field. These collaborations were deemed helpful for country partners to collectively plan the technical and financial details, exchange data and technology and formalize shared intellectual property regimes, building institutional capacities in the context. The project teams were also encouraged to participate in online research networks, discussion groups, communities and forums for collaboration and knowledge sharing. Increased number of collaborations and participation on research forums has been taken as an indicator for improvement in the research capacity of the teams.

3.1.4   Principle 4: Dissemination and Impact

Dissemination is an essential part of undertaking research. Research is as credible as much as it is referenced, cited in other publications, brought to people knowledge and properly disseminated. Dissemination of research, through peer reviewed publications and presentations at academic conferences, is essential for sharing knowledge (Harris 2004; Breen et al 2004). Capacity building for wider research dissemination incorporates instruments of publicity through factsheets, the media and the Internet (Cooke, 2005) for a variety of stakeholders, including public, policy makers and the relevant research community. The following indicators were thus used to measure the research capacity building in terms of capacity for dissemination and impact.

1. Development of a local project website
2. Organization of awareness seminars
3. Creation of promotional materials
4. Participation of teams in events and competitions

Country teams were required to provide local content for the multilingual project website maintained at www.pan110.net. In addition the teams were encouraged to host their separate websites providing detailed information about their respective groups. This was deemed necessary for providing global access to project outputs.

Country teams were also encouraged to organize awareness seminars to disseminate and publicize research results to local community. Through these seminars partner institutions could present the research work to the key stakeholders from government, IT industry, academia, media and end user communities. In addition teams have been provided funding support for developing promotional materials

including flyers, CDs containing the project outputs for marketing purpose. Country teams have also been encouraged to participate in events and competitions to present project outputs at both national and international forums. These activities have been regarded useful for improving the research dissemination capacity of the country teams.

3.1.5 Principle 5: Infrastructure Development

Rhee and Riggins (2007) defines infrastructure as a set of structures and processes that are set up to enable the smooth and effective running of research projects. These include availability of technical resources including equipment, books, connectivity, etc. as well as sound academic and managerial leadership and support for developing and sustaining research capacity. Thus availability of requisite infrastructure is necessary for improving upon the research capacity. In this context, following indicators were used to measure the advancement in infrastructure development to conduct research:

1. Acquisition of academic resources
2. Procurement of equipments
3. Provision for Operating expenses

The project had planned to build the appropriate localization research infrastructure in the country teams by providing funds for acquiring academic resources, e.g. books and journals, specialized software and (to some countries with) support to cover the recurring administrative expenses of the organizations. These indicators were thus measured to evaluate the extent to which the project had contributed to improve upon the infrastructure development in the country team.

3.1.6 Principle 6: Sustainability and Continuity

Wignaraja (2009) defines capacity development as a process of transformation that emerges from within the individuals, organizations and systems. Long term sustainable capacity development requires consolidation of local systems and processes through practice. Effective research capacity building therefore must establish strategies for maintenance and continuity of organizational skills development and maturity of processes and procedures for similar future endeavors.

Indicators defined to measure sustainability and continuity of research through the project are defined below:

1. Degree of organizational skill development
2. Number of resources trained in different domains of localization

PAN Localization project was designed to create multi-disciplinary teams including technical developers, linguistics, social scientist, policy relevant personnel, etc. for effective project execution. This was to ensure the sustainability and continuity of research even after the project's conclusion.

3.2 Data Collection Methodology

Mixed Method approach was adopted to collect data regarding the indicators of principles of research capacity building. Initially structured questionnaire (attached in appendix A) was sent through e-mail to each of the relevant country project coordinators for getting information on capacity building. In addition, supporting information was also acquired from the annual progress reports and research reports respectively developed by the regional secretariat of the project and the partner institutions. After data collection and analysis, the level of research capacity building conducted through the project has been assessed and described in the country findings sections below in perspective of the ICT landscape in the country component.

3.3 Country Findings

3.3.1 Afghanistan

In Afghanistan PAN Localization project has been working in collaboration with ACSA, http://www.acsa.org.af/ and MCIT, http://mcit.gov.af/en.

Computer literacy in Afghanistan was estimated at less than 10 percent of the population in 2010 (Bureau of Democracy, Human Rights & Labor, 2010). This underdevelopment has been mainly attributed to the political environment of the country that has been prevailing during last three decades. When ICT was introduced to the world, the Afghan society was in the process of reshaping itself after the war. The resultant digital divide became one of the major problems of the social sector of Afghanistan. To overcome this problem, quick development in information technology sector has been required. For quick development, it was necessary that the technology should be understandable; and content should be available in the most commonly spoken local language of the country, so that Afghanis could conveniently acquire the latest technology.

Through PAN Localization project, the situation gradually improved in ICT sector of Afghanistan. Marjan (2009, p. 132) explained that *ACSA, in collaboration with the MCIT and Microsoft, completed the Pashto version of Microsoft Windows XP and Office 2003 in December 2007. Work on font, lexicon, and spell check development is still in progress. The ACSA team has likewise prepared the initial feasibility report and produced the localized version of IDNs in the Pashto language. All of these measures are expected to boost the capacity of the Afghan people to develop the digital content*.

The following section presents data on each of the six principles of research capacity building impacted through the project.

3.3.1.1 Skill Development

The fundamental focus of the project was to build indigenous research skills in local language computing so that country teams are able to develop localized software. The development of these localized software involved expertise in linguistics, computer science and computational linguistics. During the first phase of the project, Afghanistan country component had worked in the areas of character set finalization, Pashto font and keyboard layout development.

Figure 3.1: Trainees of Afghanistan, 2007

As the team was newly formed and had little expertise in the domain, therefore a five day training on <u>Localization Essentials</u>, was conducted at the regional secretariat to jump start the team. This training was focused on the topics related to localization requirements, font development, keyboard, collation, lexicon, XML framework, spell checker and advanced localization applications. This training was very helpful in providing knowledge of basic localization topics to the Afghanistan team, as most of the team did not have any prior experience in local language technology development.

During phase II, the deliverables required from the team included Pashto Sea Monkey and Pashto character set for IDNs. Thus to impact relevant knowledge, training was organized at the regional secretariat focusing on different steps involved in font development, OS software localization, and Outcome Mapping. The training was very helpful both technically as well as in instigating evaluative thinking, by providing an orientation to planning, monitoring and evaluation through outcome mapping methodology. After receiving this training, the team updated Pashto keyboard for Windows XP and Vista along with Pashto fonts. In addition, the team was able to develop Pashto character set for IDNs and Pashto Sea Monkey comprising of web browser and email client.

Afghanistan				
Deliverables Required	Ling.	CL	CS	Status
Phase I				
Pashto Font	*		*	Completed
Pashto Keyboard			*	Completed
Phase II				
Pashto Sea Monkey	*		*	Completed
Pashto Character Set for IDNs	*		*	Completed

Table 3.1: Afghanistan Team's Status regarding Localized Software Development

The following table 3.1 summarizes the status of software deliverables required during the project's Phase 1 and 2, along with the respective skills required in Linguistics (Ling), Computational Linguistics (CL) or Computer Sciences (CS) to develop the respective software.

As presented in the table above, the knowledge of computer science and linguistics were required for the development of Pashto Font, Pashto Sea monkey and Pashto Character Set for IDNs while development

of Pashto keyboard involved expertise in computer science only. Afghanistan team successfully accomplished these research outputs. It is also worth mentioning that the country project team has also advanced in their over-all local language computing skill set through continued research and development A comparison of the accomplished localized software during project's phase 1 and phase 2 reveal the fact that the project country component was researching on intermediate complexity language computing software like Pashto Sea Monkey as compared to Phase 1, in which, the team was only focusing on the developments of basic localization software like Pashto Font, Pashto Keyboard etc. The team also had the honor of winning the Manthan award (2008) in the E-Localization category for Pashto SeaMonkey. Thus it is evident that the team significantly gained technical skills during the project implementation. Furthermore, Afghanistan country component developed a glossary of around 3000 terms in local language to ease and standardize the translation of Microsoft LIP Localization and this performance manifested skill capacity enhancement of their team.

Afghanistan country component's performance regarding localized software was significant in spite of the fact that their project team could not manage publication of relevant research papers.

3.3.1.2 Development of Linkages

To execute the project, ACSA collaborated with the MCIT at the national level. This collaboration was very effective as MCIT was the organization responsible to address ICT issues in the country and this greatly helped to start the localization project in the country. At the international level, ACSA collaborated with Microsoft Localization Project through which Microsoft assisted the Ministry in further development of localization project. These collaborations enabled the partners to collectively work out the technical and financial details, exchange data and technology and discuss and formalize shared intellectual property regimes, building institutional capacities in the context.

3.3.1.3 Dissemination

The Afghanistan country team hosted their separate website www.acsa.org.af/ providing detailed information about their research groups, hosted by their organizations, which are linked from the main PAN L10n project's website as well. This has given global access to the project outputs.
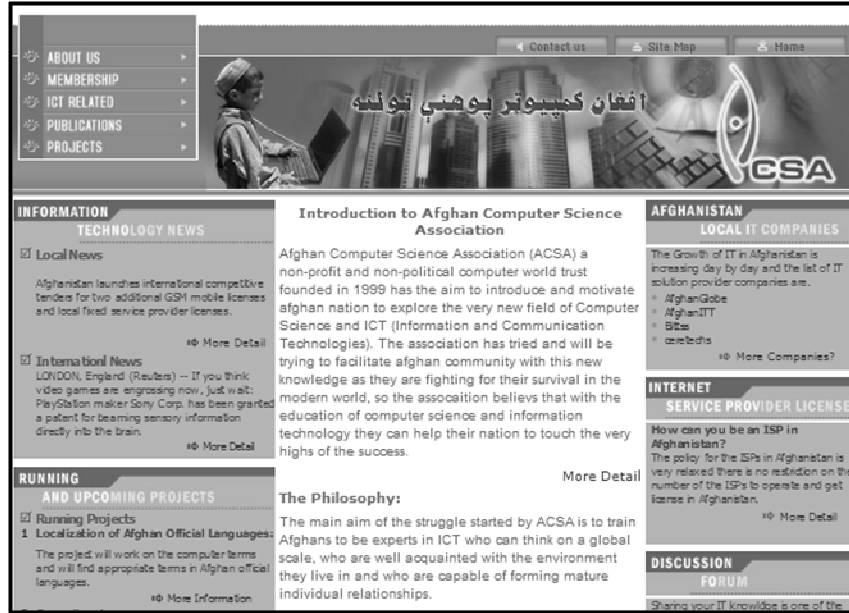
Figure 3.2: Home Page of Local website of ACSA

In addition through the project, ACSA has organized awareness seminars to disseminate and publicize research results to local community. The first ever National Computational Linguistics Seminar in Afghanistan was conducted on August 12th-13th, 2006 by the country team. It was the first step to promote awareness of local language computing standards on such a large scale. The objective of this seminar was to provide a platform for networking between the linguists and ICT professionals.

Figure 3.3: Awareness Seminar 2006

The work done by Afghanistan country component of PAN Localization project was also highlighted in the seminar by the country leader Mr. Omar Mansoor Ansari. This seminar contributed towards evoking public awareness about the project and its research work. Moreover, it was a quick and stable move towards advancement of the localization program by meeting with high authorities of the Government of Afghanistan. It was an attempt to bridge the gap between linguistic scholars and ICT professional by

creating a joint committee from the members of these two sectors. During phase II the Afghanistan component participated in Manthan Award and ICT Mela South Asia 2008. The team won the award in the E-Localization category for Pashto SeaMonkey.

Figure 3.4: Mr. Rafiqullah Kakar is receiving Manthan Award South Asia 2008

3.3.1.4 Infrastructure Development

During both the phases of the project, the country team was provided funding to support monthly stipends for the researchers, acquisition of equipment and books related to different disciplines like linguistics, language processing and computer science. Equipments included PCs, scanners and printers. The availability of funding to support these expenses aided in the development of appropriate localization research infrastructure and enhanced research capacity in Afghanistan.

3.3.1.5 Sustainability and Continuity

In terms of sustainability of research, the organizational research capacity was enhanced as the project team graduated from the research on standardization and basic localization during phase 1 to a step ahead research in intermediate complexity localized software development during the project phase 2. In addition, the project trained seven researchers in the local language computing disciple belonging from different domains like management, CS and linguistics that have later taken the work forward and acquired new projects in the localization research domain.

3.3.2   Bangladesh

In Bangladesh, the project was executed by CRBLP, http://crblp.bracu.ac.bd/ in Phase 1 and in collaboration with D.Net, http://www.dnet.org.bd/bid.htm) during Phase 2.

ICT is considered to be a key enabler in the development of a country. The usage of personal computer per 100 populations in Bangladesh was 0.02 in 1997, which rose to 0.34 in 2002 (Government of Bangladesh & United Nation, 2005) which is still quite low as compared to that of the developed world. One of the significant factors of this low usage is the lack of local content which limits the use of ICTs specially the Internet (Raihan, 2009); therefore, for rapid development in ICT sector of Bangladesh

localization is a pre-requisite. Therefore, PAN Localization project took the initiative to enable the non-English speaking people of Bangladesh to have access to the ICT means by developing of localized software. During the project's phase 2, the D.Net arranged training for end users to learn the use of the developed localized software.

The following section presents data on each of the six principles of research capacity building impacted through the project.

3.3.2.1 Skill Development

During phase I of the project, Bangladesh country component had planned to develop Bangla Lexicon, Bangla Optical Character Recognition (OCR) and Bangla Spell Checker. However in addition to the agreed deliverables for the project, the country team also developed Bangla Pad and Bangla Collator during the phase I.

For the development of Bangla OCR system, at the start the HMM model had been used for system training and recognition. This system had been matured by adding preprocessing and post processing modules. The preprocessing modules had been used to correct the input image and in the post processing module dictionary and some rules had been used to correct the recognized word. After attaching these modules, the reported accuracy was about 98%. The Bangla OCR had been integrated with open source OCR framework OCRopus. The research work on the Bangla OCR has been published in the form of three research papers titled "Integrating Bangla script recognition support in Tesseract OCR", "Rule based segmentation of lower modifiers in complex Bangla scripts" and "Elimination of splitting errors in printed Bangla scripts" and published in Conference on Language and Technology 2009 (CLT09).

During phase II of the project, Bangladesh country team worked on TTS for Bangla, Bengali SMS to Speech Application, Bengali Wordnet, Bengali Diphone Database, Bengali Speech Database, English-Bengali Parallel and Aligned Tagged Corpus, 5M Word tagged corpus, Bangla gTLD and ccTLD. In addition, Bangla Cleaned Corpus, Bangla Legal Text Corpus and Bengali Speech corpus had also been developed during phase II.

Bangladesh team had successfully released Bangla-English version of PENN Treebank parallel corpus for first 100,000 words. In addition to this work a Bangla corpus of 5 million words had also been gathered covering various domains such as scientific, medical, newspaper articles and samples from novels, stories, textbooks as well as transcribed speech. The work on part of speech tagset had also been carried out and a tagset of 55 tags had been developed. Using this tagset, 25,000 words had been manually tagged. The Brill tagger had been trained for automatic tagging of the Bangla text. The reported accuracy of this tagger was around 70.6 %. Bangla Parallel corpus had been tagged with this tagset. The Bangla country component had worked on language table and translations of gTLDs and ccTLDs in Bangla for IDNs. Major issues faced during the design of character set, gTLD and ccTLDs had also been reported. In the second phase, the team had also been working on the development of the Bangla WordNet (BWN) based on English WordNet (distribution of Princeton University). The primary focus of BWN had been on design and implementation of a framework which could be used to build and use Bengali WordNet. For the development of the BWN, the bottom up approach had been used which translate the words in the

target language. The high frequency 6,000 words from the Prothom Alo corpus had been selected for the BWN development.

The synsets had been compiled in lexical source files, which had been then included into the WordNet database using a "grinder", and the resulting system could be used through a set of interfaces. The Bangla WordNet is online available at http://www.bracuniversity.net/research/crblp/demo/bangla_wordnet/bwnV1.50/. The research paper titled "BWN - A Software Platform for Developing Bengali WordNet" has been published in International Joint Conferences on Computer, Information, and Systems Sciences, and Engineering (CISSE 08).

Figure 3.5: Bangla WordNet Page

Bangla team had developed two speech corpora; one was read speech corpus and other was diaphone corpus. To develop this speech corpus 106,860 words text corpus had been collected from different domains such as magazine, novels, blog, legal text, a small part of constitution of Bangladesh, history, and different types of news. A professional speaker had been hired for recording. After recording, the sentence level labeling is performed on the cleaned corpus. This corpus had almost 10,000 sentences and 18,000 unique tokens. This speech corpus had been used to develop acoustic models for speech recognition, to analyze the intonation pattern, and to develop a TTS by unit selection technique. In addition to this speech corpus, the diphone database had also developed. It contains 4,355 sentences, which were typically nonsense sentences. These sentences had been formed by combining the nonsense words with 4,355 diphones. Bangladesh component had been working on Text to Speech systems during first phase of project. A working Text to Speech application had been released. The system was based on Festival; an open source TTS engine. During the development of this system, various language resources had been gathered including speech corpora, pronunciation lexicon and text to sound rules. The system had been thoroughly tested for two aspects (Intelligibility and Naturalness) and three levels of results had been reported. Synthesized speech was 85% accurate for sentence level, 84% for phrases and 57% for words in terms of intelligibility. The degree of naturalness was 90% for sentences, 85 % for phrases and 57% for words. Future directions and improvements had been suggested by researchers in published paper "Text To Speech for Bangla Language using Festival", in Proc. of 1st International Conference on Digital Communications and Computer Applications (DCCA2007), Irbid, Jordan, 2007. The Bangla TTS is winner of the 2010 innovation award of BASIS.
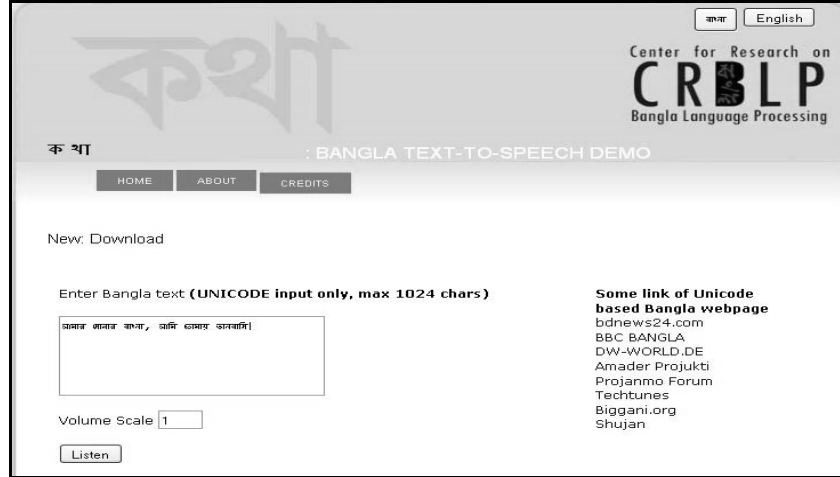
Figure 3.6: Bangla Text to Speech System

The following table summarizes the status of software deliverables required during the project's Phase 1 and 2, along with the respective skills required in Linguistics (Ling), Computational Linguistics (CL) or Computer Sciences (CS) to develop the respective software.

Bangladesh				
Localized Software	Ling.	CL	CS	Status
Required Localized Software of Phase I				
Bangla Lexicon	*		*	Completed
Bangla OCR System	*	*	*	Completed
Bangla Spell Checker	*	*	*	Completed
Other Localized Software of Phase I				
Bangla Pad	*		*	Completed
Bangla Collator	*		*	Completed
Required Localized Software of Phase II				
TTS for Bangla	*	*	*	Completed
Bengali SMS to Speech Application	*	*	*	Completed
Bengali Wordnet	*	*		Completed
Bengali Diphone Database	*	*	*	Completed
Bengali Speech Database		*	*	Completed
English-Bengali Parallel and Aligned Tagged Corpus	*		*	Completed
5M Word Tagged Corpus	*			Completed
Bangla gTLD and ccTLD	*		*	Completed
Other Localized software of Phase II				
Bangla Cleaned Corpus	*			Completed
Bangla Legal Text Corpus	*			Completed
Bengali Speech Corpus	*		*	Completed

Table 3.2: Bangladesh Team's Status regarding Development of Localized Software

Bangladesh research team had been able to deliver all the agreed software outputs. The project's research work significantly enhanced the team's capacity. This was evident from the fact that during phase two, the team was majorly involved in the advanced localization technology development. With the help of research carried out during PAN Localization project, Bangladesh team had also released Language Interface Packs for Microsoft platform.

During the project's phase 2, D.Net had been focusing on empowerment of the rural poor through its rural livelihood Information network. In this network, content developers were trained to develop locally relevant content and information workers/ Infomediaries were trained to enable the rural community to access the localized technology. D.NET organized two trainings to train Infomediaries.

First training, titled "Accessing and Intermediation of Local Language Content" was conducted by D.NET to train Infomediaries on word processing, spreadsheet Fundamentals, and basics of Word Wide Web followed by a refreshers training on similar topics. These trainings and workshops helped Bangladesh team in enhancing their skill enormously as indicated by their steady progresses in localized content development.

Figure 3.7: Infomediary Training on Local Language Content Development

Ability to publish research in the form of research papers is a salient indicator for measuring the researcher's research capacity. The project team of Bangladesh published 8 research papers covering MT, Script and Speech processing during the project's phase 2 which is the second highest number of papers published by a team within all the collaborating PAN Localization project teams. Detailed list of research report published by the project team is presented in Appendix B.

3.3.2.2 Training to Conduct Close to Practice Research

In order to establish the need for localized application, specific questions were asked from project's beneficiary communities regarding the language that they speak at home and at their work. Answers from this question would ascertain their preference of language to undertake everyday communication, both written and verbal. When end-users were asked regarding the language spoken at home and work, 100% respondent indicated that they only use local language for communication at home as well as at their workplace. The respondents were also asked to rate their reading skill and writing skill in English on a scale ranging from excellent to poor. Eleven respondents in total answered this question and 4 of them rated their reading skill in English as excellent and none of the respondents rated their reading skill in

English as poor. Three respondents rated their writing skill in English as excellent and only 1 respondents rated their writing skill in English as poor. Similarly the respondent were also asked to rate their reading skill and writing skill in the local language on a scale ranging from excellent to poor. Eleven respondents in total answered this question and 7 rated their reading skill in local language as excellent and no one of respondents rated reading skill in local language as poor. Again 7 rated their writing skill in local language as excellent and none of the respondents rated their writing skill in local language as poor. These results are presented in the graphs below.
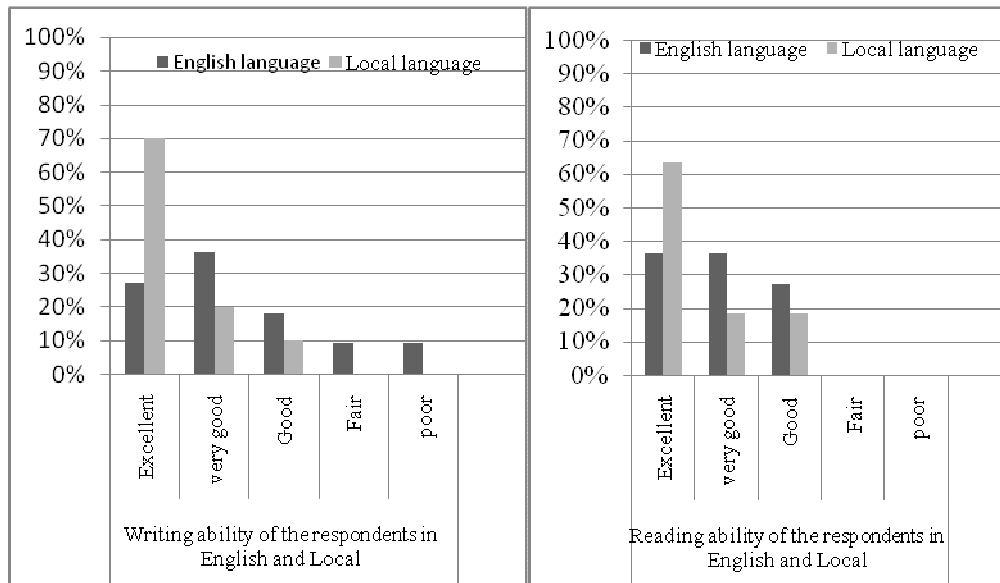
Figure 3.8: Graphs Showing Reading and Writing Ability of the Respondents of Bangladesh, in English vs. Local Language

Thus researching for development of local language ICT applications becomes directly useful and relevant to the subject communities, because in order to communicate electronically, and for work, the communities would require applications developed in local languages of the communities. These responses from the local communities of Bangladesh help establish the fact that local language computing research initiated by the project is predominantly required by the nation. It is very *close to practice* and urgently required for national development to proliferate the use of ICTs in the country.

3.3.2.3 Development of Linkages

Bangladesh was encouraged to establish partnerships and collaboration with institutions that had more expertise in a specific field. Through PAN Localization project, at national level, CRBLP collaborated with D.Net. This collaboration enabled D.Net to train rural communities of Bangladesh to use the localized software developed by CRBLP.

The project team also had been participating in online research networks, discussion groups, communities and forums for collaboration, knowledge sharing and learning. The project had created an online support network to encourage project partners to be a part of an online learning culture. Bangladesh had been participating on this forum very actively by sharing their project experiences with others. Nepal and Bangladesh team discussed their challenges in developing spell checker for open source software for

Brahmic scripts on this online network and the solution to this problem based on the use of HunSpell by Nepalese helped the Bangladesh team to develop Bangal spell checker.

3.3.2.4 Dissemination

Project team of Bangladesh also created awareness about its work through its local website http://crblp.bracu.ac.bd/. Development of this site has been instrumental in providing global access to the project outputs.

Figure 3.9: Homepage of Local Website of Bangladesh

The project team of Bangladesh also followed other different strategies to disseminate its work and create awareness of the advantages of local language technology. In training seminar on Content Development Methodologies organized by D.Net key stakeholders from NGO sector were invited and the country team provided an overview of the various activities undertaken as part of the project. Another awareness Seminar on "Computational Linguistics in the Bangla Language: Current Situation and Future Prospects", Organized by The Linguistics Association of Bangladesh, was organized at the Linguistic Department, Dhaka University.

The CRBLP at BRAC University launched the first official release of the Bangla language processing software packages for TTS and OCR named "Katha" and BanglaOCR. The activities of CRBLP were widely covered by the newspapers, IT magazines and electronic media such as The Daily Star, The New Age, the Prothom-Alo, the Monthly Computer Jagat and Bangladesh Television (BTV).

Figure 3.10: TTS Launching Seminar at BRAC University

These seminars had been very helpful in creating awareness and Bangladesh project team received e-Content and ICT for Development contest 2010 award as a finalist.

Figure 3.11: Professor Mumit Khan Receiving the Award

In addition, Bangladesh Software Industries Association (BASIS) selected Bangla TTS, developed under PAN Phased II, for BASIS IT Innovation Search Award in SOFTEXPO 2010**.** BASIS selected three finalists for this award, and showcased the projects in their 2010 expo. Interestingly, two out of three were PAN projects - Bangla OCR and Bangla TTS.

Figure 3.12: Professor Mumit Khan is receiving Award

Participants of Bangladesh project team also attended three research paper conferences.

3.3.2.5 Infrastructure Development

During both the phases of the project, the country team was provided funding to support monthly stipends for the researchers, acquisition of equipment and books related to different disciplines like linguistics, language processing and computer science. Equipments included PCs, scanners and Laser Printer and cartridges, DVD writer, USB, Amplifier preamp, Large Speaker, Small Speakers, Headphones, Switch and networking, switch and networking. The availability of funding to support these expenses aided in the development of appropriate localization research infrastructure and enhanced research capacity in Afghanistan.

3.3.2.6 Sustainability and Continuity

In terms of sustainability of research, the organizational research capacity was enhanced as the project team graduated from the research on standardization and development of intermediate complexity localization applications during phase 1 to conducting cutting edge research on advanced localization in Bangla during the project's phase 2. In addition, the project trained 22 researchers during both the project phases in the local language computing disciple belonging from different domains like management, CS, social sciences and linguistics that have later taken the work forward and acquired new projects in the localization research domain.

3.3.3   Bhutan

In Bhutan, PAN Localization project was executed by DIT, http://www.dit.gov.bt/. DIT is the lead department working under MoIC for the development and coordination of all ICT-related activities in the country.

Information Technology is a relatively new and embryonic industry in Bhutan (Jurmi & Wangchuk, 2010). In Bhutan, the computer usage per 100 inhabitants was 1.45 in 2002 (Government of Bangladesh & United Nation, 2005) which showed that the country was faced with an enormous digital divide when compared with similar averages from other countries in the region. Thus development was urgently needed in ICT sector of Bhutan. To boost ICT development, it was mandatory that digital content should be made available in the local language (Dzongkha) so that technology might be understandable for Bhutanese. Through PAN Localization project, localized software has been developed e.g. Dzongkha Desktop, Dzongkha keyboard and Dzongkha Linux. Jurmi and Wangchuk (2009, p. 157) highlighted that *The Dzongkha Desktop is good alternative for users who cannot read and write in English because it has an interface in Dzongkha as well as in English. Dzongkha Linux, which was funded under the PAN Localization project, was a success story in the research and development (R&D) front*. With the success of the project, Bhutan is making considerable progress in implementing numerous ICT initiatives.

The following section presents data on each of the six principles of research capacity building impacted through the project.

3.3.3.1 Skill Development

During the first phase of the project, Bhutan country team worked on the development of either basic or intermediate complexity software like Dzongkha Fonts, Keyboard and Linux Distribution with open Office. The initial challenge for the project was availability of researchers with relevant linguistic and computer science background to conduct this research. Initially, the project could only hire 1 researcher to work on the project. Thus the first training on <u>Localizing Open Office in Dzongkha</u> and Mozilla Localization was conducted by Mr. Guntupalli Karunakar soon after the one person research team was formulated. The main objective of the training was to achieve Dzongkha (Tibetan script) support in ICU, and do a Dzongkha enabled OpenOffice. The country team developed localized applications including keyboard, fonts, collation and locale and all of these utilities had been incorporated in Dzongkha Linux which was released on 2nd June, 2006. It was a Debian based Linux operating system with localization of Gnome and OpenOffice.org suite. Eventually this project created demand and interest in localization research and more researchers joined the project in the later years of the project's first phase.

As a part of this project DIT released a book on Dzongkha Computer Terms containing 5,000 phrases. Bhutan team had been working on corpus collection from different domains such as arts, religion, official documents and sports. The collected corpus contains 400,000 words (600,000 syllables). The collected texts had been sourced mainly from dictionaries, printed books, the print and broadcast media, and from relevant websites. A lexicon of approximately 23,000 unique words had been extracted out of this corpus, containing meaning, pronunciation and part of speech tag of each word. DIT had also been working on Part of Speech tagset development. Proposed Dzongkha tagset contains 41 tags which are thoroughly defined with examples. Selection of tags, for this tagset, is carried out by complying PENN Treebank guidelines. The major issue faced during development of automatic tagging system was absence of word segmentation utility. In order to solve this problem 20,000 words had been segmented and tagged manually. After that Tree Tagger was trained with the help of this manually annotated sub corpus. The test results showed that automatic tagging can be done with an accuracy of nearly 85%. Future directions of this work included incorporation of word segmentation module and increase in the size of trained data.

During phase II of the project, Bhutan country component worked on the development of basic, intermediate and advanced complexity local language computing applications like on Dzongkha gTLDs and ccTLDs, Dzongkha Speech Corpus, Dzongkha TTS, Dzongkha Diphone Database, Dzongkha Corpus, Dzongkha Linux 3.0 and Dzongkha Lexicon. Language tables and lists of gTLDs and ccTLDs had been released as final output. In addition to development of these lists, a comprehensive testing exercise of IDNs had also been carried out in DIT. A lab had been set up with domain name server, web server and client machines. The test exercise had been done on Local Area Network. Client machines had been provided with Mozilla Firefox plugin for local language domain name translation to punycode. The prototype successfully worked for all modules of IDNs process. Research on Text to Speech System (TTS) had also been started by DIT in Phase I. Dzongkha phonetic set description and diphone inventory had been developed. The completed milestones after Phase I is text to diphone conversion module. In Phase II Initial work had been carried out on process of speech synthesis. The work was done in collaboration with HLT team at NECTEC. A complete model of TTS had been proposed, using Hidden Morkov Models, as an output of this activity. Unfortunately due to unavailability of skilled technical resources the work faced progress issues. In order to resolve this hurdle and enhance ICT capacity, two

team members had been sent to NECTEC, Thailand. Both of them stayed there for two months and received training on complete developmental cycle of TTS. A working prototype of TTS system was available now in which input string was converted into diphones and then speech was synthesized. Improvement of this system is continued in DIT after Phase II. Bhutan team initiated work on OCR system development during first year of Phase II. A team member from Bhutan went to attend regional training of OCR at NECTEC in July, 2007 with support from PAN Localization project. It helped in analyzing Tibetan script and a technique of ligature segmentation was proposed. Research on segmentation process had been continued after that training. Due to lack of image processing and machine learning knowledge, productive outcome for segmentation and recognition modules was not enough to cope with script challenges. Eventually, another training of Bhutan team had been arranged in 2009, in Lahore to provide the appropriate mentoring. The training was helpful in development of prototypical OCR system for Dzongkha Jomolhari font. The work on this project was being carried forward at DIT after completion of Phase II.

In the second phase most of the work done on Dzongkha Linux had been enhancement of the localization of existing Open Source Software. The work had been done in terminology translation of Gnome Interface, OpenOffice.org, FireFox, ThunderBird, Debian Installer and CD burning applications, etc. These improvements had been carried out according to the feedback of users of Dzongkha Linux. Also, the reported bugs had been removed for new release e.g. in previous versions broadband modem was not supported, which had been incorporated in DzongkhaLinux 3.0. The live CD has been release in 2010.
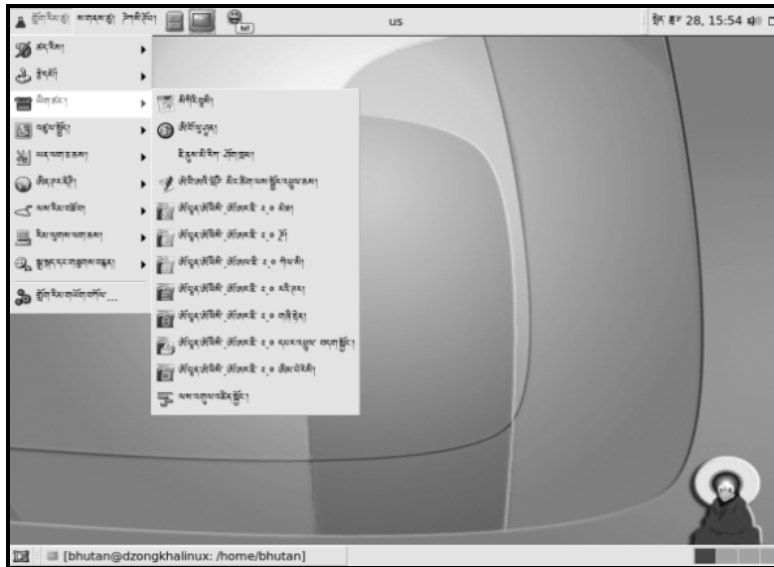
Figure 3.13: DzongkhaLinux Desktop

The following table summarizes the status of software deliverables required during the project's Phase 1 and 2, along with the respective skills required to develop the respective software.

Bhutan				
Localized Software	Ling.	CL	CS	Status
Phase I				

29

Keyboard in Linux			*	Completed
Linux Distribution with Open	*		*	Completed
Phase II				
Dzongkha gTLDs and ccTLDs	*			Completed
Dzongkha Speech Corpus		*	*	Not
Dzongkha TTS	*	*	*	Completed
Dzongkha Diphone Database		*	*	Completed
Dzongkha Corpus	*		*	Completed
Dzongkha Linux 3.0			*	Completed
Dzongkha Lexicon		*		Completed

Table 3.3: Bhutan Team's Status regarding Localized Software

The above table showed that Bhutan country team had accomplished the larger objective of the required localized software and has submitted 99% of the localized software required by the project country component as per the contract.

Based on the project experience, the country project leaders were asked to prorate the ability and skill development/enhancement of the organization's researchers during project phase 2 on a scale of 1-5; where 1 represent *Challenged*; 2 represent *Fair*; 3 represent *Average*; while 4 represents *Good* and 5 represents *Excellent* enhancement in the team's performance. The following table presents those comparative figures for assessment of project team's capacity by the project leader from Bhutan country component collected for the team's performance at the beginning of project Phase 2 in 2007 and towards the end of this phase in 2009.

Research Domain	Start of Project, Early 2007	Towards Project End, Mid 2009
LLC Project Development	2	4
LLC Project Design	2	4
Problem Identification	2	4
Project Implementation	2	4
Ability to Do Analysis	2	4
Ability to Communicate Results	3	5
Multi Disciplinary Research	1	3
Quantitative Analytical Skills	2	5
Qualitative Analytical Skills	2	4

Table 3.4: Bhutan Team's Performance regarding Skill Development

The table shown above presents that the country project leader confirmed the enhancement of team's skills starting from project development and design to its implementation and analysis within the 3 year span of the project..

The publications of research paper produced by PAN Localization project teams at various national as well as international research conferences was used as the second indicator for analyzing research capacity enhancement. Research team in Bhutan produced 1 research paper covering TTS. Detailed list of research reports publication by each country team is given in Appendix B.

3.3.3.2 Development of Linkages

To develop the technical terminology translations for the software, DIT collaborated with Dzongkha Development Authority, National Language Development Authority and Language Standardization Authority at national level .The advantage of that collaboration was that once the terminology was developed by DDA, it would become a national standard for such terminology translation. DIT collaborated with NECTEC, Thailand at international level. The Pan Localization work paved the way for a MoU between DIT Bhutan and NECTEC Thailand to promote R&D in the area of ICTs. The MoU included a plan to strengthen Open Source Natural Language Processing, Image Processing Technology and Speech Processing Technology in Bhutan. The project work also resulted in development of a permanent research and development division at DIT. The Bhutan project team also had been participating in online research networks, discussion groups, communities and forums to access to technical support.

3.3.3.3 Dissemination

Bhutan country component, hosted their separate websites http://www.dit.gov.bt/ providing detailed information about their respective research groups, hosted by their organizations, which are linked from the main website as well. This has given global access to project outputs.

Figure 3.14: Homepage of Bhutan Local Website

The team also supported the development of Bhutan Digital Library Project http://www.e-bhutan.net.bt/ndlb. The mission of the Digital Library project was to preserve and promote the cultural heritage. The website contains large volume of content in Dzongkha. The team had supported the development of a Centenary website http://www.bhutan2008.bt/ which contains the content about 100 years of Monarchy, centenary events and general information about culture, political system and constitution of Bhutan. The Bhutan component had also created a CIC web portal, http://dzongkha.sourceforge.net comprising information in various areas including health, education, government services and agriculture etc.

Figure 3.15: Homepage of www.bhutan2008.bt

The project had been organizing awareness seminars to disseminate and publicize research results to local community. These seminars had been attended by a large number of participants from academia, public and private sectors. Awareness Seminars and end user trainings had been organized by Bhutan country component to disseminate their work on localized software.

The launching ceremony of DzongkhaLinux held on June 2nd, 2006 received vast media coverage. In addition, two training sessions were conducted for end-users. The first training on DzongkhaLinux was organized in DIT, MoIC, Thimphu, Bhutan from November 19th to November 27th, 2007. The objective of the training was to train the end user trainers and it was expected from trainers that they go back to their own organizations (Private IT institutes and Government offices) and train the other users.

Second Dzongkha Linux End User training was conducted from December 3-16, 2008 at Dzongkha Language Institute (DLI), Thimphu and Computer Management Institute (CMI), Phuentsholing. Both institutes were authorized by the Dzongkha Development Commission (DDC) and DIT for conducting trainings. About 40 participants from various agencies including Royal Bhutan Army and Royal Bhutan Police were participated in these trainings. The objective of this training was to provide basic knowledge about installation and operation of Dzongkha Linux.

Figure 3.16: End Users Training (December 3-16, 2008)

Participants were also trained on open office applications, Firefox and Thunderbird and other applications. Participants were expected to further transfer their knowledge and skills to their colleagues and create awareness about Dzongkha Linux and open source software in general. Training manuals on Dzongkha Linux Installation, Open Office and other applications were provided to the participants. Trainees were also provided lessons materials in the form of PowerPoint Presentations.

A two day's training session on Dzongkha Linux was conducted in Lhuntse DYT hall on 28th May, 2009 for the teachers, monks and staffs Dzongkhag Administration. The training was funded by Dzongkha Development Commission (DDC). DIT conducted an advance level Dzongkha Linux training from 11th to 25 January 2010. 500 CDs of Dzongkha Linux 3.0 and key boards were also distributed by the country team. These trainings and distribution of CDs had been helpful to disseminate Bhutan country team's work on localized software.

3.3.3.4 Infrastructure Development

During both the phases of the project, the country team was provided funding to support monthly stipends for the researchers, acquisition of equipment and books related to different disciplines like linguistics, language processing and computer science. Equipments included PCs, scanners, printers, servers, UPS, Amplifier, Mics, speakers, CDs. The availability of funding to support these expenses aided in the development of appropriate localization research infrastructure and enhanced research capacity in Afghanistan.

3.3.3.5 Sustainability and Continuity

In terms of sustainability of research, the organizational research capacity was enhanced as the project team graduated from the research on standardization and basic localization during phase 1 to a advanced research in script processing during the project phase 2. In addition, the project trained 22 researchers in the local language computing disciple belonging from different domains like management, CS and linguistics. This has further strengthened the organizational capability in maturing their research capacity technically, and able to commit and acquire more projects.

3.3.4   Cambodia

In Cambodia, PAN Localization project has been conducted by the Committee for Standardization of Khmer Scripts in computer of the council of Ministries during phase I. Further, in Cambodia, the PAN Localization project also collaborated with NiDA, http://www.nida.gov.kh/), Institute of Technology, Cambodia and MoEYS, http://www.moeys.gov.kh/ during phase II.

According to NiDA, "Khmer is the official language of Cambodia where 90 percent of the population, about 6 million people, speaks it as the first or second language". It depicts that the most spoken language in the country is Khmer rather than English. Sorasak & Konsona(2009, p. 173) highlighted that "Cambodia does not have English language skill". English language competence issue leads to underdevelopment of ICT sector in the country because of English being the lingua franca for ICTs. Thus for the development in ICT sector, it was necessary to develop localized software so technology should be understandable for the local people of Cambodia and they feel convenient to use technology.

Through PAN Localization project, MoEYS and NiDA effectively contributed towards local language computing policy in the country. "In 2005, the Ministry of Education, Youth, and Sport (MoEYS) started to implement the policy and strategies on Information and Communication Technology in Cambodia" (Sorasak & Konsona, 2009, p.167). Ministry of Interior (MoI), Ministry of Finance(MoF), Ministry of Transportation, Ministry of Commerce (MoC) and National Assembly (NA) adopted the Khmer Unicode and the applications developed by the project teams. NiDA also regularly invited government officials in its end users trainings to create awareness among policy makers of the benefits of localized technology. All these activities show that Government of Cambodia highly interested regarding development in ICT sector. "Over the last 10 years, the government has been proactive in the development of ICT" (Sorasak & Konsona, 2009, 167).

The following section presents data on each of the six principles of research capacity building impacted through the project.

3.3.4.1 Skill Development

During phase I of the project, Cambodia country component worked on the development of localized software including Encoding conversion utility, Sorting utility for Khmer, Terminology Translation for Khmer Computer Interface, Khmer Lexicon, Khmer spell check. When the project was conceived, regional secretariat organized a six month long mentor placement in Cambodia to start the project because the available management did not have relevant experience in the domain. Thus the mentor helped identify and hire the appropriate HR, initiated project planning and execution strategy, provided relevant training, thus set-up the complete process and research framework to execute the project. This Mentor placement program on Localization and Khmer Language Processing was conducted from 20th June 2004 till 27th December 2004.

The team received training on basic to advanced programming and basic to advanced language processing techniques. The discussion during training covered the topics regarding Open Type font development, Unicode language Processing, Lexicon Development, Project life cycle from design to execution and testing, Advance programming in C++ and Visual Basic.NET. The second training was held from May 2006 - June 2007 in Cambodia that focused on development of Khmer FOSS applications.

Figure 3.17: Training on Localization and Khmer Language Processing 2004

In addition, other localized software including, Khmer spell check XP, Khmer spell check Vista Khmer spell check Vista, Khmer line breaker, Khmer Conversion, Khmer Collation and sorting, Khmer Unicode Standardization and Khmer Word-Wrap Utility had also been released by Cambodia country component in phase I. Text processing utilities, developed in phase I, were proved to be very useful for end users. It was decided to make them available across both the platforms i.e. Windows and Open source (OS). Hence, very extensive work had been done in order to release those utilities with multiple interfaces. All of these utilities were ported to Java platform and their Linux and Windows versions were release. In addition to that, these applications had been incorporated in OpenOffice.org writer application. After covering the diversity of end users, Windows based APIs were also developed and released to facilitate programmers working in Khmer Language Processing area. In addition to that Khmer Collation support for MySQL had also been provided.

During phase II of the project, Translation of gTLDs and ccTLDs in Khmer, English-Khmer Parallel and Aligned Tagged Corpus 100k words, SMS in Khmer Application, Khmer TTS, Khmer Diphone Database and Khmer Speech Corpus were required as per the contract. In phase II, the trainings were also conducted to enhance the team's technical skills. First training was conducted in Institute of Technology of Cambodia from May 16th, 2007 to June 16th, 2007, to introduce the fundamental concepts of Khmer Language Processing and the Usage of Khmer Unicode in the Web Site Development.

During phase II, Cambodia country component also focused on conducting research on advanced NLP areas like OCR, TTS, IDNs and PoS Tagging. They have also reviewed HTML standards for English and Khmer. The initial work on the IDNs started in Phase 2. This work contributes the research reports for defining the character set and encoding constraints for IDNs in Khmer. Comprehensive lists of gTLDs and ccTLDs have been released as output of IDN project. Part of speech tagset design and documentation has been carried out by PLC team. It has also gathered a corpus of 150,000 words tagged with their parts of speech. This corpus is annotated with the help of probabilistic TnT tagger. A trained model of TnT tagger was released along with corpus. The project of Khmer-English parallel corpus was planned but unavailability of appropriate translators was a hurdle for its completion. A number of interviews were carried out, by country project leader, to hire translators but the quality of translations was not good enough to kick off this project. Another alternative was to hire professional companies to do the work but there were two major issues in this approach. First, there was not any capacity building incentive involved and secondly the cost of that outsourcing was going way beyond the budget. Eventually this project was closed and budget was reallocated for training activities with consent of regional project leader.

During second year of phase II MoEYS has been working on Khmer OCR. Due to resource retention issues the progress of project was quite slow. In order to resolve this issue, the Mentor placement program for Optical Character Recognition and Khmer Language Processing was held in PAN Cambodia Office from 1st July 2008 till 10th January, 2009. The objective of the program was to introduce Cambodia team with basic local language computing technologies. The training covered topics on Khmer OCR; Open office plug-ins of Khmer applications, Khmer collation support for MySQL, Automatic POS tagger for Khmer, Khmer Lexicon development, Khmer Tagged Corpus, Khmer SMS software for Java based mobile phones. Based on the guidance of the mentor, functional OCR of Limon R1 & S1 fonts had been released http://panl10n.net/english/OutputsCambodia2.htm. The reported accuracy of recognition process is nearly 98% for ligature level. The incorporated modules in this system are Noise removal,

Preprocessing, Recognition and Unicode mapping. Following screen shot elaborated the output of Khmer OCR system.
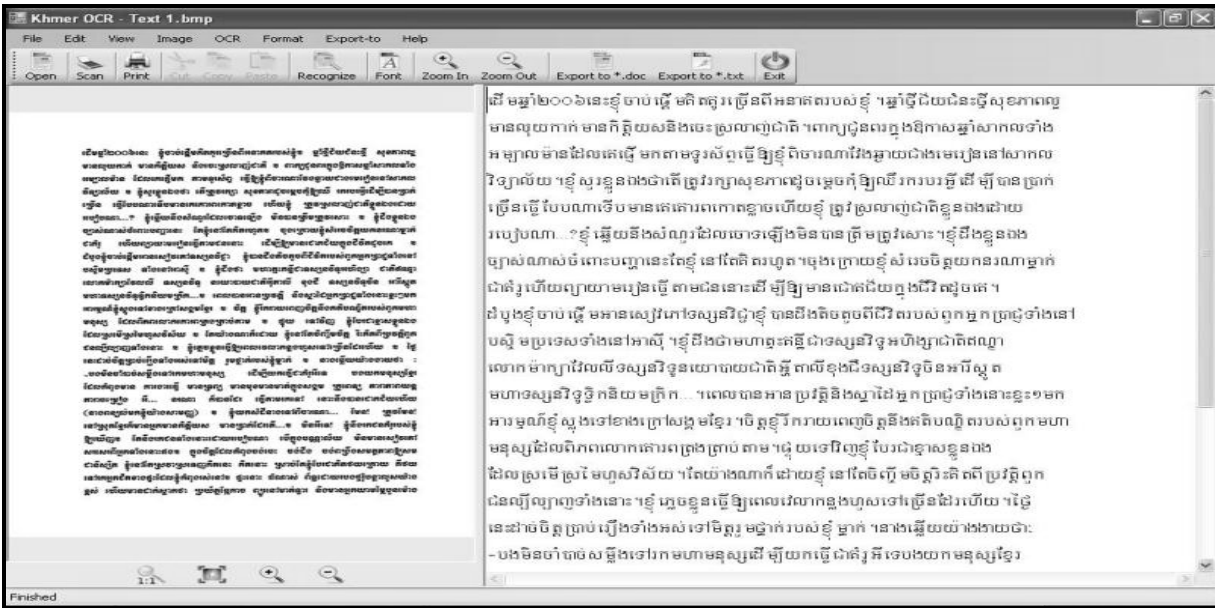
Figure 3.19: Khmer OCR

PLC team also worked on development of Khmer mobile SMS application. Research on mobile font development, keyboard design and J2ME framework had been carried out. This activity was planned for one year duration and a complete SMS application in Khmer language was released along with mobile font. This activity was helpful in exploring the domains of font rendering on mobile platform. Another bright aspect was availability of Khmer font for other mobile applications. Some small but effective sub projects had also been carried out at PLC including Khmer Typing Game, One Click Installation Package and active PAN Cambodia website http://www.pancambodia.info/index.php.

The MoEYS component had a second team of ITC as well which carried out very extensive research in speech processing area. A complete Text to Speech system had been released as an outcome of phase II. The successful development of this system had been possible because of research on Phonetic and Phonological Analysis of Khmer, NLP engine and Letter to Sound Conversion. The recordings for this system were carried out by professional Khmer speaker in isolated studio environment. This activity was good initiative for development of speech processing area in Khmer. It helped not only in gathering language resources in the area of speech but also provided an opportunity to establish a speech lab in ITC. A long term output is formation of research environment and origination of ongoing NLP research activities. Highlighted issue faced in this project was unavailability of language resources e.g. Diphone database and programming skills for language processing.

The following table summarizes the status of software deliverables required during the project's Phase 1 and 2, along with the respective skills required to develop the respective software.

Phase I				
Localized Software	Ling.	CL	CS	Status
Required Localized Software				
Encoding Conversion Utility			*	Completed
Sorting Utility for Khmer	*		*	Completed
Terminology Translation for	*		*	Completed
Khmer Lexicon		*		Completed
Khmer Spell Checker		*	*	Completed
Other Localized Software				
Khmer Spell Checker Vista		*	*	Completed
Khmer Spell Checker XP		*	*	Completed
Khmer Line Breaker	*		*	Completed
Khmer Conversion			*	Completed
Khmer Collation and Sorting	*		*	Completed
Khmer Unicode Standardization			*	Completed
Khmer Word-Wrap Utility	*		*	Completed

Table 3.5: Cambodia Team's Status regarding Localized Software in Phase I

The below mentioned table showed that Cambodia country component had been able to deliver all localized software as per the contract. Development of other outputs has proved Cambodia team's skill enhancement during the project.

Phase II				
Localized Software	Ling.	CL	CS	Status
Required Localized Software				
Translation of gTLDs and ccTLDs in Khmer	*		*	Completed
English-Khmer Parallel and Aligned Tagged	*		*	Completed
SMS in Khmer Application	*		*	Completed
Khmer TTS	*	*	*	Completed
Khmer Diphone Database	*		*	Completed
Khmer Speech Corpus	*		*	Completed
Other Localized Software				
Khmer Text to Speech System	*	*	*	Completed
Khmer OCR	*	*	*	Completed
Khmer Encoding Conversion (Java Application)	*		*	Completed
Khmer Word Segmentation/Line Breaking (Java Application)		*	*	Completed
Khmer Collation (Java Application)	*		*	Completed
Khmer Spell Checking (Java Application)		*	*	Completed
Khmer Find & Replace (Java Application)			*	Completed
Khmer SMS J2ME Application	*		*	Completed
Open Office Writer Plugin	*		*	Completed

				*	Completed
PLC Typing Tutor					

Table 3.6: Cambodia Team's Status regarding Localized Software in Phase II

Based on the project experience, the Cambodia country component was asked to prorate the ability and skill development/enhancement of the organization's researchers during project phase 2 on a scale of 1-5; where 1 represent *Challenged*; 2 represent *Fair*; 3 represent *Average*; while 4 represents *Good* and 5 represents *Excellent* enhancement in the team's performance. The following table presents those figures for assessment of its team's capacity by the project leader from country component collected for the team's performance at the beginning of project Phase 2 in 2007 and towards the end of this phase in 2009.

Research Domain	**Start of Project, Early 2007**	**Towards project End, Mid 2009**
LLC Project Development	1	4
LLC Project Design	1	4
Problem Identification	1	4
Project Implementation	1	4
Ability to Do Analysis	1	4
Ability to Communicate Results	1	4
Multi Disciplinary Research	1	4
Quantitative Analytical Skills	1	4
Qualitative Analytical Skills	1	4

Table 3.7: Cambodia Team's Performance regarding Skill Development

The table shown above presents that the project leader had confirmed the enhancement of team's skills starting from project development and design to its implementation and analysis within the 3 year span of the project.

3.3.4.2 Training to Conduct Close to Practice Research

Training on localized technology was conducted by NiDA. The training program focused on a diverse group of trainees including students, teachers, farmers, journalists, ministry officials and other staff from government organizations. Rural communities were trained on SuSE Linux, Open Office, use of Internet and email using NiDA Khmer Standard Unicode Keyboard. The training was conducted in eight different provinces of the country.

In order to establish the need for localized application, specific question was asked from the communities regarding the language that they speak at home and at their work. Answers from this question would ascertain their preference of language to undertake everyday communication, both written and verbal. When end-users were asked regarding the language spoken at home and work, 100% respondent indicated that they only use local language for communicate at home as well as at their workplace. This response clearly indicated that the language most convenient for communication for the specific communities was their respective local language.

The respondents were also asked to rate their reading skill and writing skill in English on a scale ranging from Excellent to poor. 150 respondents in total answered this question and only 2 rated their reading

skill in English as excellent and 17 respondents rated their reading skill in English as poor. Only 3 of them rated their writing skill in English as excellent and 27 of them respondents rated their writing skill in English as poor. Similarly the respondent were also asked to rate their reading skill and writing skill in Local Language on a scale ranging from Excellent to Poor. 150 respondents in total answered this question and majority of them (138) rated their reading skill in local language as excellent and no one of respondents rated reading skill in local language as poor. A large majority of them (137) rated their writing skill in local language as excellent and only 1 respondent rated writing skill in local language as poor. These responses are presented in the graphs below.
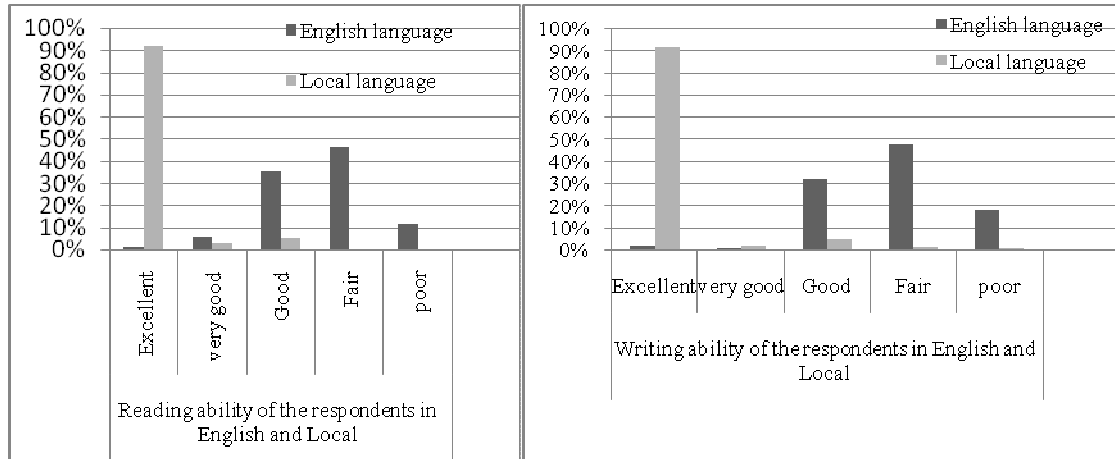
Figure 3.20: Graphs Showing Reading and Writing ability of the Respondents of Cambodia, in English vs. Local language

Thus researching for development of local language ICT applications becomes directly useful and relevant to the subject communities, because in order to communicate electronically, and for work, the communities would require applications developed in local languages of the communities. These responses from Cambodian end-users helped establish the fact that local language computing research initiated by the project is predominantly required by the nation. It is very close to practice and urgently required for national development to proliferate the use of ICTs in the country.

3.3.4.3 Development of Linkages

Through PAN Localization project, Cambodia country component collaborated with National Institute of Language and Institute of Technology, Cambodia at national level. Institute of Technology, Cambodia (ITC) that had professors working on localization research and students were taking up localization research projects in their BS final year projects. This partnership was helpful in providing a continuous pool of trained researchers to work on the localization projects. The Cambodia project team also had been participating in online research networks, discussion groups, communities and forums to access the technical support and to enhance the quality of their work regarding localization.

3.3.4.4 Dissemination

Cambodia Country component disseminated its work through its website. This has given global access to project outputs.

Figure 3. 21: Home Page of Local Website of Cambodia

In Cambodia, the project followed different strategies to market and disseminate the research. Project partners created awareness of the advantages of Khmer Unicode and Khmer based applications among the computer users and developers. Cambodia country component of PAN Localization project conducted a seminar on 23rd December, 2005 in Cambodia, to introduce the work done under the project. The seminar was attended by a large number of participants from teacher and student communities. CD containing the research outputs (Khmer Smart Typing, Encoding Conversion Utilities, Collation and Sorting Utilities, Word Wrapping Utilities, Spell Checker Utilities) was also freely distributed. The seminar was also followed by technical session. NiDA also invited government officials in its end users trainings to promote awareness of localized technology among stakeholders and general public. In his visits to Cambodia, Project Leader also created awareness among stakeholders and sensitize policy makers about the value of research being done under the project.

3.3.4.5 Infrastructure Development

During both the phases of the project, the country team was provided funding to support monthly stipends for the researchers, acquisition of equipment and books related to different disciplines like linguistics, language processing and computer science. Equipments included PCs, printers, servers, notebook and networking, Amplifier, Speaker, Mics, Speakers, Headphones and networking. In both phases of PAN Localization project, Cambodia country component focused on development of networking and conducting trainings and available funds were mostly used for these activities. The availability of funding to support these expenses aided in the development of appropriate localization research infrastructure and enhanced research capacity.

3.3.4.6 Sustainability and Continuity

In terms of sustainability of research, the organizational research capacity was enhanced as the project team starting from scratch in localization research during the project's phase 1 initiated research on advanced localization research in language, script and speech processing during the project phase 2. Through the PAN Localization project a significant number of technical developers, linguists and social

scientists have been trained. The Cambodia country component trained 62 trainers which is the highest among the all participant countries.

3.3.5   Indonesia

In Indonesia, PAN Localization project worked with BPPT, http://www.bppt.go.id and University of Indonesia (UoI), http://www.ui.ac.id/. The PAN Localization project in Indonesia was initiated in April 2008 during phase II. Through PAN Localization project, Indonesia country component further enhanced the research in local language computing. The results given below show that significant progress was made in the ICT sector of the country during the project. PAN Localization project also boosted Indonesia project team to build research capacity in local language computing. The following section presents data on each of the six principles of research capacity building impacted through the project.

3.3.5.1 Skill Development

POS Tagger for Bhasa Indonesia, 1 Million PoS Tagged Corpus, and SMT System were the localized software required to be developed through the PAN Localization project. In addition, the project aimed to develop linguistic resources and English to Bahasa Indonesia MT System. There were two corpora developed during the project. Initially BPPT collected 500,000 words of Bahasa Indonesia in order to build a corpus. This collection had been translated sentence by sentence into English. Along with this activity, UoI team translated first 500,000 words of PENN Treebank into Bahasa Indonesia in collaboration with the Faculty of Arts, University of Indonesia. Finally, by combining these two corpora a parallel corpus of one million words has been made which is annotated with part of speech.

The UoI team also worked on development of POS Tagset for Bahasa Indonesia. They have developed a tagset containing 37 POS tags. A statistical POS tagger has been developed using Maximum Entropy. The tagger was trained on approximately 150,000 manually tagged words and used for automatic tagging of one million words corpus. SMT was the technique based on sentence-level aligned parallel corpora. The team had used open source SMT tools including Giza and Moses to achieve translations. The system had been trained on one million words corpus. Bleu Scores were computed for evaluation of MT system and achieved accuracy of 0.938 for ENG-IND translation and 0.926 for IND-ENG. Details of training and testing were given in the report released with outputs http://panl10n.net/english/OutputsIndonesia2.htm. The online system is available for general public use http://translator.iptek.net.id/PANL/. The screen in Figure 3.22 below shows the input sentence from headlines on BBC Indonesia on 30th Jan. 2010 and its corresponding English translation through this online translation tool (which translates in both directions).

Figure 3.22: Online SMT for Bahasa Indonesia

The following table summarizes the status of software deliverables required during the project along with the skills required to develop the respective software.

The below status showed that the project country component had been able to deliver all localized software as per the contract.

Indonesia				
Localized Software	Ling.	CL	CS	Status
Required Localized Software				
POS Tagger for Indonesia	*		*	Completed
1 Million POS Tagged Corpus	*		*	Completed
SMT System	*	*	*	Completed
Other Localized Software				
500,000 Word Bahasa Indonesia Parallel Corpus with Penn Treebank	*	*	*	Completed
500,000 Word Bahasa Indonesia Corpus and Parallel English Translation	*	*	*	Completed

Table 3.8: Indonesia Team's Status regarding Localized Software

Based on the project experience, the country project leader was asked to prorate the ability and skill development/enhancement of the organization's researchers during the project on a scale of 1-5; where 1 represent *Challenged*; 2 represent *Fair*; 3 represent *Average*; while 4 represents *Good* and 5 represents *Excellent* enhancement in the team's performance. The following table presents those comparative figures for assessment of project team's capacity by the project leader from Indonesia country component

collected for the team's performance at the beginning of project Phase 2 in 2007 and towards the end of this phase in 2009.

Indonesia				
	BPPT		UoI	
Research Domain	Start of Project, Early 2007	Towards Project End, Mid 2009	Start of Project, Early 2007	Towards Project End, Mid 2009
LLC Project Development	4	4	5	5
LLC Project Design	4	4	5	5
Problem Identification	4	4	5	5
Project Implementation	4	4	5	5
Ability to Do Analysis	4	4	5	5
Ability to Communicate	4	4	5	5
Multi Disciplinary Research	4	4	4	4
Quantitative Analytical Skills	4	4	5	5
Qualitative Analytical Skills	4	4	5	5

Table 3.9: Indonesia team's Performance Regarding Skill Development

The publication of research papers produced by PAN Localization project teams at various national as well as international research conferences was used as the second indicator for analyzing research capacity enhancement. The project team of Indonesia published 2 research papers covering MT, SLP, POS during the project's phase 2. Detailed list of research report publication by project team of Indonesia is presented in Appendix B.

3.3.5.2 Development of Linkages

At national level, BPPT collaborated with UoI. The collaboration had been helpful in term of the initiation of research on ICT accessibility program, Indonesia POS tagger and statistical machine translation system.

At international level, UoI collaborated with Tech University of Vienna and Kyoto University in developing local language resources and tools. BPPT also had been participating in online research networks to access the technical support in enhancing the quality of their work regarding localization and also to acquire latest information on local language computing development. The online participation had been helpful in developing speech to speech translation system.

3.3.5.3 Dissemination

Indonesia country component hosted their separate websites http://www.bppt.go.id/ providing detailed information about their respective research groups, hosted by their organizations, which are linked from the main website as well thus providing global access to project outputs.

Figure 3.23: Home Page of Local Website of BPPT

BPPT organized awareness seminar in Jakarta on August 11, 2008. This seminar was attended by Lecturer and students of University of Indonesia and Atmajaya Catholic University. Dr. Sarmad from Pakistan also attended this seminar. University of Indonesia organized two Malindo workshops on 2nd August 2009 in Singapore and on 2nd August of 2010 in Jakarta.

To disseminate research outputs, Dr. Hammam Riza, CPI leader of BPPT had also presented the paper Building Parallel Text Corpora for Multi-Domain Translation System at 7th Workshop on Asian Language Resource, ACL-IJCNLP Singapore in August 2009. The local website organization of workshops helped Indonesia to disseminate their research on localization. The team from UoI developed FASILKOM Web Services (FWS) to disseminate their research outputs http://fws.cs.ui.ac.id/SampleClientPOSTag/index.jsp. The POSTAG application was accessible through this web service (see Figure 3.23). The research paper on Indonesian POS tagger was presented by Dr. Mirna Adriani of Indonesia component (head of Information Retrieval Lab: http://ir.cs.ui.ac.id at the Third International MALINDO Workshop, ACL IJCNLP 2009 in Singapore and the Fourth International MALINDO Workshop 2010.

3.3.5.4 Infrastructure Development

During the project, the country team was provided funding to support monthly stipends for the researchers, acquisition of equipment and books related to different disciplines like linguistics, language processing and computer science. Equipments included computer hardware like PCs, scanners, printers and server. Indonesia country component also focused on development of networking and available funds were mostly used for these activities. The accessibility of these funds helped developing appropriate localization research infrastructure and enhanced research capacity in Indonesia.

3.3.5.5 Sustainability and Continuity

Through PAN Localization project 18 technical developers, linguists and social scientists have been trained to enable sustainability and continuity of the research being undertaken. The availability of funding to support these expenses aided in the development of appropriate localization research infrastructure and enhanced research capacity in the country.

3.3.6   Laos

In Laos, the project was executed by NAST, http://www.laol10n.info.la/, the policy making body for ICTs in the country. Thus the work done under the project had direct influence on local language computing policy in the country.

In Laos, the majority (73 percent) of the country's 5.65 million people live in the rural areas and the usage of computer per 100 habitants is just 1.7 in 2007 (Phissamay, 2009). A fundamental reason for the low ICT diffusion in the respective country is due to the English language competence issue because only a few people in Laos understand English and all available software was in English language. Although, the Government of Laos has been encouraging the officials and students to learn English, yet, the official and dominant language in Laos is Lao (Bureau of East Asian and Pacific Affairs, 2011). "Government of Laos aims to bring the country into information age by increasing general access to ICT" (Phissamay, 2009, p.243. Firstly, in order to facilitate access of common Lao people to ICT, it was needed that software should be developed in local language so that majority of the people in Laos could easily use the latest technology.

Through PAN Localization project, Lao country team made localized software and paved the way for significant and fast development in ICT sector of the country. The PAN Localization project encouraged the country team to build research capacity in local language computing as well. The following sections present the impact of the project in building research capacity in Laos as studied using the six principles of RCB discussed earlier.

3.3.6.1 Skill Development

During phase I of the project, Laos country component worked on Sorting Utility for Lao, Spell Checker, Lao Lexicon and Grammar Checker. In addition, Laos country component also worked on proposals for standards development for Lao fonts, LaoPad and Keyboard during the first phase of the project. Various utilities such as open type fonts, encoding conversion, collation sequence and word segmentation were developed with the guidance of regional mentor http://panl10n.net/english/activity 11.htm, through the training on Localization in Lao, conducted in Laos by Mr. Aamir Wali, from 4th – 11th September 2004. The main topics covered during training were Keyboard Layout creation, Collation and Line/word breaking, basic concepts in C++, Open Type Font Development. The training had been organized to enables Laos country team to develop algorithm for line/word breaking and collation. Furthermore, training on Computing for Localization, conducted at STEA, Laos by Mr. Nadir Durrani, from 27th January 2005 till 27th June 2005. The main objective of this training was to capacity build the Laos team in local language computing and enables them to localize software components in Lao.

The training covered topics including basic programming and algorithm development, Advanced Concepts in C++, Visual Basic DOT NET, Microsoft Visual Studio, Using Microsoft Access with VB.NET, Development of Line Breaking Algorithm, Development of Collation Sequence and Sorting Algorithm, Encoding conversion of Non-Unicode fonts to Unicode fonts, Development of Lao-English-Lao Lexicon, Development of Find & Replace Utility and Development of Spell Checker.

Figure 3.25: Training on Computing for Localization in 2005

During phase II of the project, Lao Keypad Standard, Lao Natural Language Processor for TTS, List of gTLDs and ccTLDs in Lao, English-Lao Parallel and Aligned Tagged Corpus 100k words, Lao Open Office with Lao Line Breaking and Collation, Phonetic module for Lao TTS, Lao speech corpus, Lao TTS, Lao diphone database and 100 mb Lao content development were required localized software to be developed by the country component.   In this context, Laos country component had worked on the development of the Lao-English Parallel corpus of PENN Treebank. The team had successfully released approximately 50,000 words corpus.   The research had also been done for the Part of Speech of Lao language. The tagset of 41 tags had been developed. The tagging of corpus could not be carried out because of lack of linguistic knowledge. In addition, the lexicon of around 67,000 words had been developed from different dictionary books and non electronic sources. The Laos team had worked on language table and terminology translation of gTLDs and ccTLDs in Lao for IDNs. The Lao translated gTLDs and ccTLDs can be accessed online http://www.panl10n.net/english/OutputsLaos2.htm. In addition, Laos country component also worked on Lao OCR, Corpus Analysis Tools, Corpus Analysis Tools, Openoffice.org plug-in for Lao Line breaking and collation and Microsoft Office plug-in for Lao Line Breaking and Collation during phase II. Lao OCR had been initially developed in the first phase with the support of Sri Lanka team.   Later on, it had been enhanced by adapting the algorithm developed by the NECTEC for Thai OCR. The neural network had been used for the pattern matching. This newer version of Lao OCR had provided the facility of input image of multiple formats and it generates the respective Unicode of the characters.   This OCR could recognize printed Lao characters of ten mostly used fonts namely Alice0Lao, Alice1_2000, Alice2_2000, Chanthabuli Lao, Chanthabuli 95, aysettha Lao, Saysettha 2000, Saysettha 95 and FontLao1. The reported accuracy of the system was around 98.7%.

Laos country team had also started research on TTS. Work on phoneme analysis of Lao language, which was a precursor for the development of TTS application has been completed. Though much progress had been made in terms of application development, it had been largely achieved through the mentor placement program for porting applications and with the support of NECTEC for OCR and TTS. The

country level agreement between Laos team and the Thai team at NECTEC will certainly contribute to this end after completion of PAN L10n project.
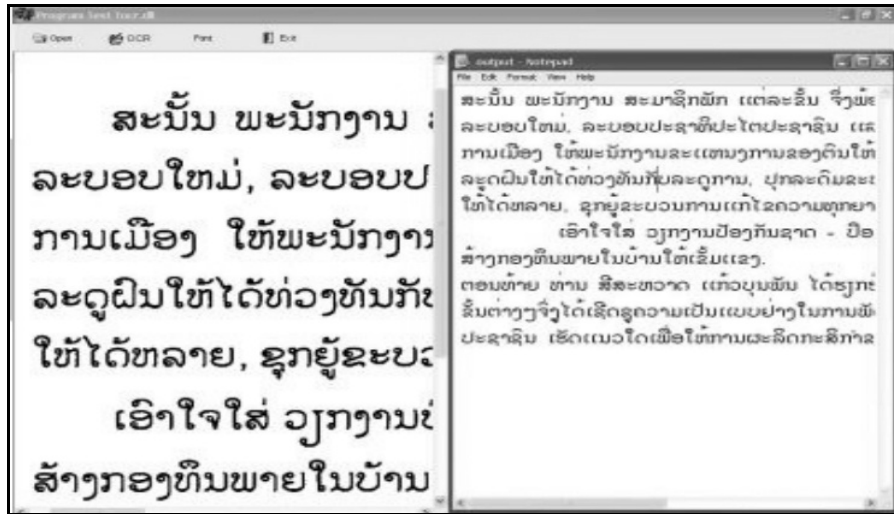
Figure 3.26 Lao OCR

The process of development of Plug-ins for the Microsoft Office and OpenOffice.org writer had two modules. In the first module, the algorithms for Encoding Standardization and Lao Syllabification had been used which were implemented with the help of APIs of all these utilities developed in phase 1. The second module was embedding of this work which had been developed to extract the data from the office application. After applying the languages dependent algorithm, data was also saved back into the respective places. This module had been developed with the help of Cambodian component by customizing the Cambodian code for Lao language. In addition to this work OpenOffice.org interface Localization had also been started and currently 7,000 words have been translated. The work will be continued at STEA (NAST) after closure of PAN Localization process. All of the above mentioned research outputs are available online at http://www.panl10n.net/english/OutputsLaos2.htm.
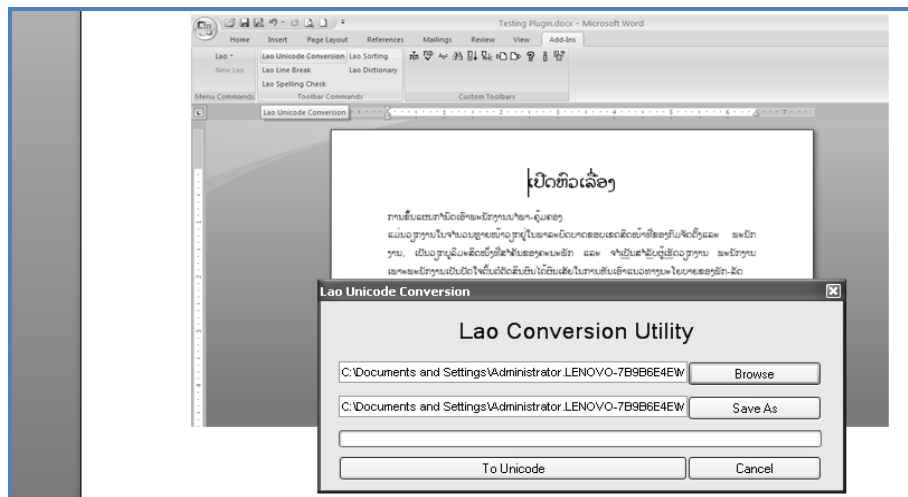
Figure 3.27: Lao OpenOffice.org Plug-in

Evaluation Findings of PAN Localization Project

The following table summarizes the status of software deliverables required during the project's Phase 1 and 2, along with the skills required to develop the respective software.

Laos				
Localized Software	Ling.	CL	CS	Status
Required Localized Software of Phase I				
Sorting Utility for Lao	*		*	Completed
Spell Checker		*	*	Completed
Lao Lexicon		*	*	Completed
Grammar Checker		*	*	Not completed
Other Localized Software				
Lao Font				Completed
Lao Pad				Completed
Lao Keyboard				Completed
Required Localized Software of Phase II				
Lao Keypad Standard	*		*	Not Completed
Lao Natural Language Processor for TTS	*	*	*	Not Completed
List of gTLDs and ccTLDs in Lao	*		*	Completed
English-Lao Parallel and Aligned Tagged Corpus 100k words	*		*	Completed
Lao Open Office with Lao Line Breaking and Collation	*		*	Completed
Phonetic Module for Lao TTS	*	*	*	Not Completed
Lao Speech Corpus	*		*	Not Completed
Lao TTS	*	*	*	Not Completed
Lao Diphone Database	*		*	Not Completed
100 Mb Lao Content Development			*	Not Completed
Other Localized Software				
Lao Sea Monkey	*		*	Completed
Lao OCR	*	*	*	Completed
Corpus Analysis Tools			*	Completed
Openoffice.org Plug-in for Lao Line Breaking and Collation	*		*	Completed
Microsoft Office Plug-in for Lao Line Breaking and Collation				Completed

Table 3.10: Laos Team's Status regarding Localized Software

Laos however has not been able to submit the required localized software due to the unavailability of trainers and the country project team has been able to submit only 40 % software as per the contract.

3.3.6.2 Dissemination

Laos country component had hosted their separate website http://www.laol10n.info.la/ providing detailed information about their respective research groups, which are also linked from the centrally maintained, PAN Localization project's website. This has given global access to project outputs.

48

Figure 3.28: Website of Lao L10n

The second website was developed for local community. This website was produced by using Dokuwiki web content development tool and contains 100 MB of content including text and images (http://laocontent.info.la). The website contains various types of information: Public Health, Information Technology, Agriculture, Law, Environment and PAN Lao localization. The content focused on the community with low literacy level or basic computer/internet skills.

NAST organized a meeting of policy makers and IT people on January 10, 2008. In this meeting PAN Localization project team in Laos highlighted the research work done under the project and raised the awareness of decision makers and other relevant stakeholders about the potential and impact of local language computing. The president of NAST appreciated the efforts being done under the project and showed his commitment to adopt a vigorous localization policy. NAST also decided to work for E application applying the research of PAN Localization team and to seek government approval for national standard on localization. In addition to disseminate work, the Laos country team also organized an ICT fair and prime minister of Laos was invited for inauguration. The team conducted a technical workshop and 70-80 government officials, students and journalists also participated in this workshop. CD containing fonts, keyboard utilities and open office application was also distributed to disseminate work on localization.
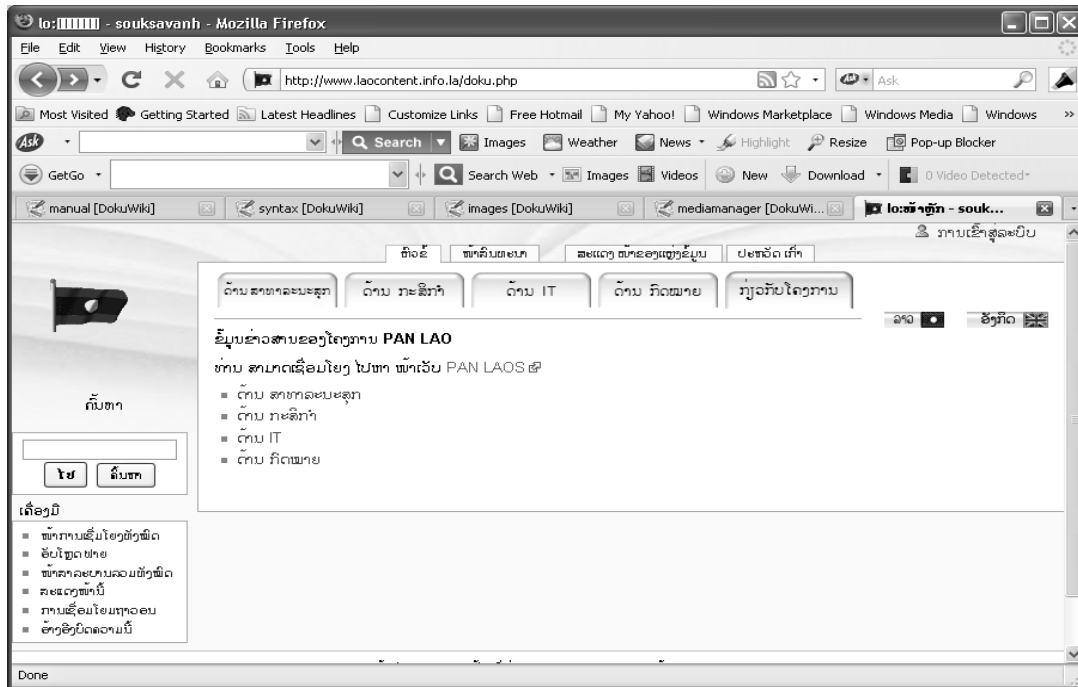
Figure 3.29: Homepage Laos Local Website of http://laocontent.info.la

End user trainings had been helpful in dissemination of research work. Laos team of PAN Localization project also planned to conduct two types of end user trainings i.e. content providers training and local community training. The country team decided to train 20 to 30 content providers on Open office, Email and Mozilla Firefox and content provider would be selected from different organizations such as newspapers, ministries such as agriculture, health, culture. Local community would be trained on email and how to copy files to CD. The team also developed a user guide on how to access, upload and download content from web, which is freely available at http://www.dokuwiki.org/manual/.

3.3.6.3 Infrastructure Development

During both the phases of the project, the country team was provided funding to support monthly stipends for the researchers, acquisition of equipment and books related to different disciplines like linguistics, language processing and computer science. Equipments included PCs laptop, switch, PC Upgrades, amplifier, headphones, mics, speaker, LCD, digital camera whereas software included Visual Studio.Net . In the operational field, participants regarding different domains of the PAN Localization project were trained and available funds were mostly used for these activities. The availability of funding to support these expenses aided in the development of appropriate localization research infrastructure and enhanced research capacity in the country.

3.3.6.4 Sustainability and Continuity

Laos country component focused on the development of basic localization and script processing during PAN Localization project. Through PAN Localization project a significant number of technical developers, linguists and social scientists have been trained to enable sustainability and continuity of the

research being undertaken. Lao country component trained 23 participants from different domains like management, technology and linguistics. However the project could not sustain the research as the largest number of technical training has been provided to Laos, but still this has not helped a lot in building the institutional capacity to advance research in this area.

3.3.7   Mongolia

In Mongolia, PAN Localization project was conducted by Mongolian University of Science and Technology, Mongolia (http://www.must.edu.mn/beta3/), National University of Mongolia, (http://old.num.edu.mn/) and InfoCon Co. Ltd www.infocon.mn. The PAN Localization project in Mongolia was initiated during phase II. The following section presents data on each of the six principles of research capacity building impacted through the project.

3.3.7.1 Skill Development

Specific localized software including list of gTLDs and ccTLDs in Mongolian, Localized FireFox (seamonkey) Web Browser, Localized Thunderbird or other Email Client, 100K Word Manually Tagged Corpus, 5 Million Word Mongolian Corpus, Tagged 10k Word Lexicon, Mongolian Spell Checker, Localized Open Office Word Processor, Localized Open Office Spread Sheet, Mongolian POS Tagger, 3K Word ASR for Mongolian, Localized Open Office Presentation Software, Localized GAIM Chatting Program (Pidgin), Integrated OO in Mongolian, Mongolian Sentence Parser, Mongolian Prototype Computational Grammar, Large Vocabulary ASR for Mongolian and Mongolian Speech Corpus were required to be developed through the PAN Localization project. In addition, Mongolia country component also worked on Localized Chat Client, Localized Seamonkey Web Browser and email client, POS Tagger and corpus builder and 3k Word ASR for Mongolia.

NUM had been working on Mongolian corpus collection from different sources such as literature, law and news paper. Total of 5 million words corpus had been collected and cleaned using the corpus cleaning tools. In addition, some spelling mistakes were corrected with the help of spell checker. The corpus is available at http://panl10n.net/english/OutputsMongolia2.htm. The NUM team had also developed 10,000 word lexicon from this corpus and is online available for downloading. Mongolian team NUM had also worked for the development of Mongolian spell checker. The main focus of developing this tool was to correct the spelling mistakes in the corpus of 5 million words. For this, a dictionary based spell checker was developed. The Figure given below shows screen shot of Mongolian spell checker.
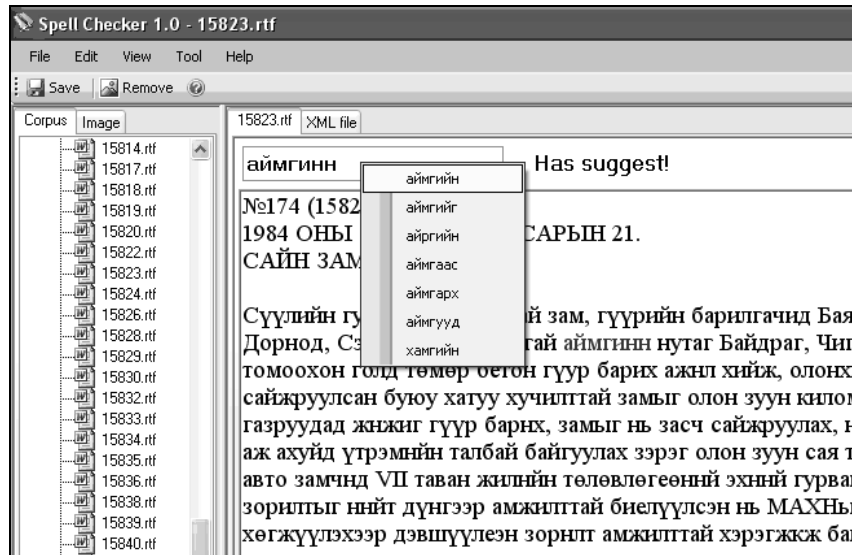
Figure 3.30: Working View of Spell-Checker

NUM had also been working on Part of Speech tagset. The tagset of 81 tags had been developed. The 100,000 words corpus was tagged manually and was online available. To manually tag the corpus, the tool for manual tagger was also developed. The statistical tagger was developed for automatic tagging of Mongolian text which was based on the HMM model. The reported accuracy was this system was around 81 % for 260,000 words. The paper title "Part of speech Tagging for Mongolian Corpus" was published in 7th Workshop on Asian Language Resources, (ACL-IJNLP2009). The Mongolia Country Component had already worked on terminology translation through involvement of a committee defined through ICT Agency of Mongolia. During the first year of second phase of the project, the Infocon team had worked on translations of gTLDs and ccTLDs in Mongolian.

The Mongolian country component had also worked on the localization of tools. The InfoCon team had carried out the localization of the web browser, email client and chat tools. A complete version of SeaMonkey internet suite which consists of email client, composer, html editor, chat, web browser had been localized with 43,000 strings. The terminology translation of chat program (PIDGIN) had also been done and released. This software is downloadable from the project's website. MUST started research for the development of an ASR. At start of the research, team conducted surveys for comparative evaluation of ASR toolkits. They had developed a prototype recognition system for Mongolian on HTK and Sphinx toolkits. The HMM based approach had been used to recognize the speech file.

The team had developed a 6,000 word ASR based on HMM by using HTK toolkit. The system had been enhanced by increasing the size of speech corpus. Total of 6,000 high frequency words had been selected from the corpus. One wave file had 10 words and each word has been repeated 10 times. Therefore the speech corpus has 60,000 wave files. The wave file had sampling rate of 16 kHz and sampling size of 16 bits. For the recording of these words 80 speakers had been selected. The test data consists of 100 words spoken by 20 native speakers who had not participated in the training data. The reported accuracy of the system was above 90% for this test data. The paper titled "A Large Vocabulary Speech recognition System for Mongolian language" was published in proceedings of Oriental-COCOSDA 2008. All of the outputs discussed above are available online at http://panl10n.net/english/OutputsMongolia2.htm. The

country team worked on IT terminology translation and the work was approved by ICTA committee. The country team worked on IDNs standards and reviewed and released generic TLDs (gTLDs), country-code TLDs (ccTLDs) for Mongolian. The Pan Localization project also supported the development of Speech Lab at MUST and Center for Language Processing at NUM with a view to ensuring the sustainability of localization work. The project also contributed to realization of an MOU signed between NUM and NECTEC.

The following table summarizes the status of software deliverables required during the project, along with the respective skills required in linguistics, CL or CS to develop the respective software.

Mongolia				
Localized Software	Ling	CL	CS	Status
Required Localized Software				
List of gTLDs and ccTLDs in	*		*	Completed
Localized FireFox (Seamonkey) Web	*		*	Completed
Localized Thunderbird or other Email	*		*	Completed
100K Word Manually Tagged Corpus	*		*	Completed
5 Million Word Mongolian Corpus	*		*	Completed
Tagged 10k Word Lexicon	*		*	Completed
Mongolian Spell Checker	*	*	*	Completed
Localized Open Office Word Processor	*		*	Not Completed
Localized Open Office Spread Sheet	*		*	Not Completed
Mongolian POS Tagger	*		*	Completed
3K Word ASR for Mongolian	*		*	Completed
Localized Open Office Presentation	*		*	Not Completed
Localized GAIM Chatting Program	*		*	Completed
Integrated OO in Mongolian	*		*	Not Completed
Mongolian Sentence Parser	*	*	*	Not Completed
Mongolian Prototype Computational	*	*	*	Not Completed
Large Vocabulary ASR for Mongolian	*			Completed
Mongolian Speech Corpus	*		*	Completed
Other Localized Software				
Localized Chat Client	*		*	Completed
Localized Seamonkey Web Browser	*		*	Completed
PoS Tagger and Corpus Builder		*	*	Completed
3k Word ASR for Mongolia	*	*	*	Completed

Table 3.11: Mongolia Team's Status regarding Localized Software

The above mentioned table showed that the project team of Mongolia has been able to submit 67% localized software as per contract.

Based on the project experience, the country project leaders were asked to prorate the ability and skill development/enhancement of the organization's researchers during project phase 2 on a scale of 1-5; where 1 represent *Challenged*; 2 represent *Fair*; 3 represent *Average*; while 4 represents *Good* and 5 represents *Excellent* enhancement in the team's performance. The following table presents those

comparative figures for assessment of team's capacity by the project leader from partner country collected for the team's performance at the beginning of project Phase 2 in 2007 and towards the end of this phase in 2009.

MUST, Mongolia		
Skill Development	Start of Project, Early 2007	Towards Project End, Mid 2009
LLC Project Development	3	5
LLC Project Design	3	5
Problem Identification	3	5
Project Implementation	3	5
Ability to Do Analysis	3	4
Ability to Communicate	2	4
Multi Disciplinary Research	3	4
Quantitative Analytical Skills	3	4
Qualitative Analytical Skills	3	4
NUM, Mongolia		
Skill Development	Start of Project, Early 2007	Towards Project End, Mid 2009
LLC Project Development	1	4
LLC Project Design	3	5
Problem Identification	2	5
Project Implementation	3	5
Ability to Do Analysis	3	5
Ability to Communicate	2	4
Multi Disciplinary Research	3	5
Quantitative Analytical Skills	1	3
Qualitative Analytical Skills	2	4
InfoCon, Mongolia		
Skill Development	Start of Project,	Towards Project
LLC Project Development	2	5
LLC Project Design	3	5
Problem Identification	3	5
Project Implementation	3	5
Ability to Do Analysis	2	4
Ability to Communicate	2	4
Multi Disciplinary Research	3	5
Quantitative Analytical Skills	3	4
Qualitative Analytical Skills	3	4

Table 3.12: Mongolia's Team (MUST, NUM and InfoCon) Performance regarding Skill Development

The above mentioned table showed that the country project leaders had confirmed the enhancement of team's skills starting from project development and design to its implementation and analysis within the 3 year span of the project specifically in over-all project execution.

The publications of research paper produced by PAN Localization project teams at various national as well as international research conferences was used as the second indicator for analyzing research capacity enhancement. The project team of Mongolia published 6 papers covering POS, Corpus, and Speech during the project's phase 2 which was the third highest among the all participating countries. Detailed list of research report publication by Mongolia project team is presented in Appendix B.

3.3.7.2 Development of Linkages

Project country component had been focusing on building capacity by developing appropriate linkages, partnerships and collaborations and for this purpose. MUST, NUM and InfoCon collaborated with each other at national level.

In addition, InfoCon collaborated with ICTPA and MUST collaborated with Institute of language and literature of Mongolian Academy of Science. ICTPA had been helpful in making IT related policy and implementing agency whereas Institute of language and literature of Mongolian Academy of Science helped in research regarding Linguistics.

Mongolia Country component had been interacting regularly with ICTPA. The meetings deliberated on open source software localization, standardization of terminology, IDNs and other related issues. The project team also had been participating in online research networks to enhance the quality of their work regarding localization and also to acquire latest information on Local Language Computing development.

3.3.7.3 Dissemination

Mongolia Country Component hosted their separate websites www.infocon.mn providing detailed information about their respective research groups, hosted by their organizations, which are linked from the main website as well. This has given global access to project outputs.

Figure 3.31: Website of INFOCON

The project had organized awareness seminars to disseminate and publicize research results to local community. These seminars had been attended by a large number of participants from academia, public and private sectors. Mongolia country component organized awareness seminar on 17 October 2009 and it was focused on Mongolia speech processing. Seminar had been attended by students and lecturers from Mongolia academy of science.

The research work of the project was presented by the country team at different workshops and conferences such as ALRN 2007 (Asian Language Research Network) from March 1-2, 2007, Tokyo, Japan, 7th Workshop on Asian Language Resources from August 2-7, 2009, Suntec, Singapore, Asian Language Resource Summit from March 20-21, 2009, Phuket, Thailand and Workshop on Applied NLP and Language Resource development from February 23 – 27, 2009, Bangkok, Thailand. Project partners had been involved in designing, developing and disseminating the material developed, which has contributed to mutual capacity to disseminate research.

3.3.7.4 Infrastructure Development

During the project, the country team was provided funding to support monthly stipends for the researchers, acquisition of equipment and books related to different disciplines like linguistics, language processing and computer science. Equipments included computer hardware like PCs, scanners, printers and speech equipments (amplifier, mics, etc). Mongolia country component also focused on development of networking and available funds were mostly used for these activities. The availability of funding to support these expenses aided in the development of appropriate localization research infrastructure and enhanced research capacity in the country.

3.3.7.5 Sustainability and Continuity

MUST focused on the development of language processing, speech processing and InfoCon focused on Standardization and Basic Localization during PAN Localization project. National university of Mongolia NUM focused on all LLC Domains namely Standardization, Basic Localization, Lang. Processing, Script processes and Speech Processing. Mongolia country component trained 27 participants from different domains like management, technology and linguistics. In terms of sustainability of research, the organizational research capacity was enhanced as the project team was undergoing advanced complexity localized software development during the project.

3.3.8 Nepal

In Nepal, the project was executed by Madan Purskar Pustakalya, http://madanpuraskar.org/ in phase 1 and by MPP in collaboration with the E-Networking Research and Development, http://www.enrd.org/ in phase 2.

In Nepal, the usage of computer per 100 populations was 0.35 in 2001(Government of Bangladesh & United Nation, 2005). The deep rooted nature of social exclusion in Nepal has hampered the development of ICT sector. Besides, English language competence issue is also one of the reasons for less progress in ICT sector because most spoken language in Nepal is Nepali but all available software was in English. For fast development regarding ICT, it was required that digital content should be transferred in local

language of the country. So that Nepali population could easily use software and acquire latest technology. Through PAN Localization project, the Nepal country team has successfully developed localized software e.g. NepaLinux. The impact of the PAN Localization Project in Nepal has been very positive in the local language computing. The work in itself was technically challenging because for the first time, everything was being done in the Nepali Language. Within a time span of around 11 months, the first result of this project came out in December 2005 with the release of the NepaLinux 1.0. This was a major breakthrough in the history of ICT sector of Nepal. In Nepal, PAN Localization project has also been helpful to build research capacity in local language computing.

The following section presents data on each of the six principles of research capacity building impacted through the project.

3.3.8.1 Skill Development

During the first phase of the project, Nepal made excellent progress by working on translation of terminology, locale and collation. A complete Nepali Linux Distribution was developed and released on 22nd December, 2005 as well http://www.nepalinux.org/. The launch was a huge success in terms of public response. More than 5 national newspapers did the coverage of the launching ceremony.

Parallel corpus generation for languages of PAN partner countries is one of the prime objectives of Phase II. Nepal country component has carried out this work in collaboration with the Pakistan team. The end goal was to develop a tri-lingual tagged corpus for Urdu-English-Nepali languages. This project was supported by PAN Localization project through funding of Language Resource Association (GSK) of Japan. It was very extensive activity including translation of text, research and development of part of speech tagset for Nepali and Urdu languages. Parallel teams in Nepal and Pakistan worked together to accomplish this task in time. Eventually the target of translating 100, 000 words is achieved and tagged parallel corpus had been released http://www.crulp.org/software/ling_resources.htm. MPP had been working on development of computational POS tagset for Nepali language. Initially a tagset was designed comprising of 112 tags. Results of semi automatic tagging showed that the designed tagset was error prone because of its depth. The tagset was carefully pruned and essential 43 tags had been selected. A corpus of 80,000 words had been manually tagged and TnT tagger was trained on it. The reported accuracy of this tagger is 97% for known and 56% for unknown words. This tagger was used to annotate developed corpus of one million words.

During phase II of the project, NepaLinux Training Kit, NepaLinux 3.0, Nepali Grammar Checker, Nepali Spell Checker, Reviewed List of gTLDs and ccTLDs in Nepali, Nepali Computational Grammar and Nepali OCR were requisite localized software to be delivered. Nepal Country Component (Madan Puraskar Pustakalya(MPP) in collaboration with its Country Partner Institution , E-Network Research and Development(ENRD) successfully conducted a Training from 9th to 17th November, 2008, in Nangi, Myagdi, Nepal for skill enhancement of project team. During training discussion was focused on Nepalinux Feedback, shortcomings, keyboard layouts, new spellchecker, machine translations, nepali sabdakos, tuxtype, gcompris, blog, CMS, domain registration, domain hosting, hard disk partitioning etc.

A meeting with policy makers was organized on 24th March, 2009 to discuss the challenges faced during development of Nepali IDNA. Some of these challenges were usage of IDNA for Nepali and alternative of symbols such as period "." and www. Efforts made by international and national bodies for enabling IDNA were also brought into the knowledge of policy makers. The meeting concluded on recommendations for future directions. MPP had worked on Nepali Spell Checker which was based on Hunspell open source framework. A lexicon of 37,000 words had been incorporated in Hunspell along with 1800 affix rules. The spell checker was incorporated in OpenOfffice.org and tested for approximately 2000 words. Reported accuracy for this test data was around 90%. The system had coverage of 6.2 million Nepali words. Nepali spell checker stand alone application had been developed and released http://panl10n.net/english/OutputsNepal2.htm.Nepali team had successfully released NCGA during Phase II. A rule based chunker had been developed along with 30 manually extracted chunking rules. Currently there were 11 chunk tags in chunkset. Nepali Grammar Analyzer was an integrated application that comprises POS tagger, chunker and a parser. Currently the NCGA parses and analyzed declarative sentences with only one verb. Around 700 verbs had been handled in this grammar checker. The accuracy of grammar analyzer depends on the accuracy of chunker. Reported results showed that it parses around 90% sentences accurately.

The work on Nepalinux had continued in second phase and its version 3.0 had been developed and formally released on 25th May, 2008. Recent research in advanced areas of NLP had been incorporated in this version. One of such utilities is Nepali Text-to-Speech application (not developed directly through PAN Localization projected). Nepali Sabdakos (Dictionary) had also been added which contains meaning of 8,000 words with examples. Other useful applications which were included in this version are offline English dictionary, Gcompris, TuxType, Nepali Spell Checker and KTouch typing tutor. MPP had also separately released an educational version of Nepalinux 3.0 (Educational) on 6th July, 2008. It contains basic Nepalinux and essential educational applications such as Spell Checker, Gcompris and KTouch typing tutor. Madan Puraskar Pustakalaya (MPP) had also released the Nepali Office CD. It comprises localized OpenOffice.org suite in Nepali along with a few other useful free and Open Source Software like Gimp, FireFox and Thunderbird etc.

The Nepali Spell Checker was also integrated in the OpenOffice.org suite. Work was being done in the area of Nepali OCR and a beta version had already been developed by the team. This had been done with the guidance and direct training from the Bangladesh team. Although complete OCR system was planned to release but the person working on this project had left the team. The work could not be transferred to other resources. Eventually the OCR project was closed with release of beta version. Under the initiatives of MPP and Kathmandu University (KU), efforts were continuing for developing a Tesseract based Nepali OCR. Language tables and lists of gTLDs and ccTLDs had been released during phase II. More information on the activity is available at http://nepalinux.org/index.php?option=com_content&task=view&id=46&Itemid=53.

The following table summarizes the status of software deliverables required during the project's Phase 1 and 2, along with the respective skills required in linguistics, CL or CS to develop the respective software.

Nepal				
Localized Software	Ling	CL	CS	Status
NepaLinux			*	Completed
NepaLinux Training Kit			*	Completed
NepaLinux 3.0			*	Completed
Nepali Grammar Checker	*	*	*	Completed
Nepali Spell Checker	*	*	*	Completed
Reviewed List of gTLDs and ccTLDs in Nepali	*		*	Completed
Nepali Computational Grammar	*	*	*	Completed
Nepali OCR		*	*	Completed

Table 3.13: Nepal Team's Status regarding Localized Software

The above status showed that the Nepal country component had been able to deliver all required localized software as per the contract.

Based on the project experience, the country project leaders were asked to prorate the ability and skill development/enhancement of the organization's researchers during project phase 2 on a scale of 1-5; where 1 represent *Challenged*; 2 represent *Fair*; 3 represent *Average*; while 4 represents *Good* and 5 represents *Excellent* enhancement in the team's performance. The following table presents those figures for assessment of its team's capacity by the project leader from country component collected for the team's performance at the beginning of project Phase 2 in 2007 and towards the end of this phase in 2009.

Skill Development	Start of Project, Early 2007	Towards Project End, Mid 2009
LLC Project Development	3	5
LLC Project Design	3	4
Problem Identification	3	5
Project Implementation	3	5
Ability to Do Analysis	4	5
Ability to Communicate	3	5
Multi Disciplinary Research	3	4
Quantitative Analytical Skills	3	4
Qualitative Analytical Skills	3	5

Table 3.14: Nepal Team's Performance regarding Skill Development

The table shown above presents that the project leader had confirmed the enhancement of team's skills starting from project development and design to its implementation and analysis within the 3 year span of the project specifically in over-all project execution.

Ability to publish research in the form of research papers is a salient indicator for measuring the researcher's research capacity. Thus publications of research paper produced by PAN Localization project teams at various national as well as international research conferences was used as the second

indicator for analyzing research capacity enhancement. Nepal country component published one research paper covering NLP during second phase of PAN Localization project.

3.3.8.2 Training to Conduct Close to Practice Research

In Nepal, end-user training on local language technology was conducted by ENRD. The training program focused on farmers, women, students, and youth groups. The training was conducted in five different locations of the country. End users were trained on NepaLinux and other localized software, such as content management system, etc.

In order to establish the need for localized application, specific questions were asked from the communities regarding the language that they speak at home and at their work. Answers from this question would ascertain their preference of language to undertake everyday communication, both written and verbal. When end-users were asked regarding the language spoken at home and work, 100% respondent indicated that they only use local language for communicate at home as well as at their workplace.

This response clearly indicated that the language most convenient for communication for the specific communities was their respective local language. Thus researching for development of local language ICT applications becomes directly useful and relevant to the subject communities, because in order to communicate electronically, and for work, the communities would require applications developed in local languages of the communities.

The respondents were also asked to rate their reading skill and writing skill in English on a scale ranging from Excellent to poor. 102 respondents in total answered this question and none of the respondents rated reading skill in English as excellent and 10 respondents rated their reading skill in English as poor. 8 respondents rated their writing skill in English as excellent and only 1 respondent rated their writing skill in English as poor.

Similarly the respondent were also asked to rate their reading skill and writing skill in Local Language on a scale ranging from Excellent to Poor. 102 respondents in total answered this question and none of the respondents rated reading skill in local language as excellent while 14 respondents rated their reading skill in local language as poor. Only 5 rated their writing skill in local language as excellent and 10 respondent rated writing skill in local language as poor. These results are presented in the graphs below.
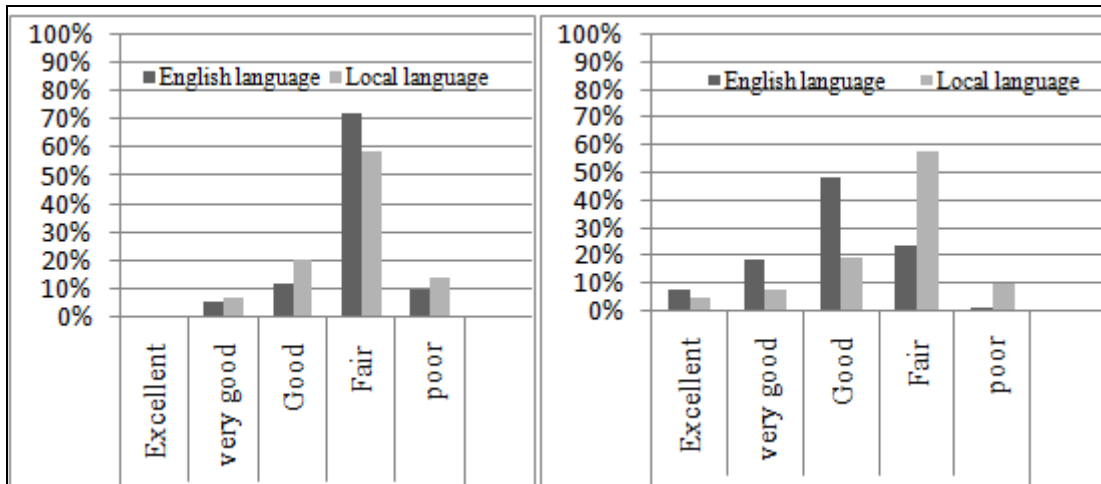
Figure 3.32: Graphs showing reading and writing ability of the respondents of Nepal, in English vs. Local language

These responses from the local communities helped establish the fact that local language computing research initiated by the project is predominantly required by the nation. It is very *close to practice* and urgently required for national development to proliferate the use of ICTs in the country.

3.3.8.3 Development of Linkages

At the national level, MPP collaborated with ENRD whereas at international level, Nepal country component collaborated with CRULP and CRBLP, BRAC University. CRULP helped Nepal country team in administration of the PAN Localization Project and provided them technical help on Language and Script Processing. CRBLP, BRAC University had been helpful for Nepal project team in developing OCR. The project created an online support network to encourage project partners to be a part of an online learning culture. The project partners had been participating on this forum, sharing their project experiences with each other. Nepal and Bangladesh team discussed their challenges in developing spell checker for open source software for Brahmic scripts. The solution based on HunSpell by Nepalese helped the team develop Bangal spell checker in Bangladesh.

3.3.8.4 Dissemination

MPP and ENRD developed their local websites accessible from http://www.mpp.org.np/pannepal/ and http://www.enrd.org/panproject/ respectively to disseminate their work.

Figure 3.33: Website of ENRD
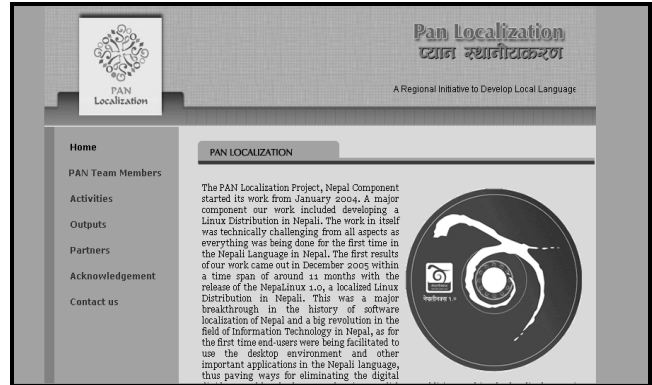
Figure 3.34: Website of MPP

Through the project, MPP also organized awareness seminars to disseminate and publicize research results to local community. These seminars had been attended by a large number of participants from academia, public and private sectors. Through these seminars partner institutions have been regularly presenting their work to the key stakeholders from government, IT industry, academia, media, and end user communities. The project had organized awareness seminars to disseminate and publicize research results to local community.

Nepalinux had been launched by Nepal country component of PAN Localization project, at Kathmandu, Nepal on 23rd December, 2005. To announce the release Mr. Kamal Mani Dixit, president of Madan Puraskar Pustakalaya (MPP) presented a complete CD of NepaLinux to Mr. Sanjeev Rajbhandri, the chief of ISP and IT Company Mercantile Communications. Speaking at this launch Mr. Sanjeev said that the Nepali operating system would gain popularity once the government and private institutions endorsed and public embraced it.

Figure 3.35: Launching Ceremony of Nepalinux

NepaLinux, had been awarded the prestigious international APC Chris Nicol FOSS Prize 2007, jointly with Free Geek (an organization based in the United States of America and working for the promotion of FOSS), The APC Chris Nicol FOSS Prize was a biennial prize established to honor a long time FOSS advocate and activist as well as APC member.

NepaLinux had been put in the annual exhibitions like CAN Info Tech organized by the Computer Association of Nepal (CAN) for the last four years continuously. An estimated 3000 copies of CDs/DVDs of NepaLinux both the downloadable and CD/DVD burnt versions had been distributed to the end users. NepaLinux is currently deployed in around 10 telecenters (Fulchowki, Dhading, Sindhupalchowk, Myagdi, Kaski, Rasuwa, Dailekh and other places - established partly under the PAN Localization Project(http://panl10n.net), Bhasha Sanchar Project(http://bhashasanchar.org) and other collaborations)and in the process of deployment in another 16 telecenters ;Bhaktapur and Butwal and 14 others under the Rato Bangla Public Private Partnership Network and facilitated by Madan Puraskar Pustakalaya in direct partnership with Nepal Telecommunication Authority. More detailed information is available at http://nepalinux.org.

In addition to dissemination of the research outputs, rural community in Nepal was trained on localized technology in five different locations of the country. The locations included Danda Gaun in Rasuwa district, Jhuwani in Chitwan district, Tolka in Kaski district, Nangi and Shika in Myagdi district. The training program focused on farmers, women, students, youth and other groups. Training program adopted the train-the-trainer format and training was conducted in three stages.

At first stage, in November 2007, MPP, in collaboration with ENRD trained telecenter operators and teachers in 10 day trainer's training program. The main objective of the training was to make the participants familiar with NepaLinux and other localized software such as Content Management System so that they can successfully provide technical support and further training to the local community. These participants were further trained in 9 day trainer's training program in home village of Magsaysay Award winner - Mr. Mahabir Pun. At second stage, every trained teacher/ tele-center operator developed his/her own training outline and course and nominated a group of five participants. Each group has representation from target population of women, farmers, students, youth and teachers. Telecentre operator/teacher trained his/her group. The trainees at this level were defined as local level leadership. At third stage, each local level leader trained his/her own community members. Trainees at third layer were identified as end-user community. 25 end-users were trained at each location and total 125 end-users were trained in that process. ENRD also evaluated the effectiveness of these trainings.

A large majority of the end users were not familiar with computer before the training and at the completion of training they were to create content. ENRD also observed that teachers and students had relatively more learning capacity and their content requirements were easier to address. ENRD also conducted further trainings on basic computer skills, open office, Net Meeting, Instant messenger and web browser (Mozilla Firefox).

Figure 3.36: Second TOT Training at Nangi, Nepal

Telecenter operators and teachers trained at first stage of training also developed five websites (www.shikha.com.np, www.nangi.com.np, www.jhuwani.com.np, www.tolka.com.np, www.dandagaun.com.np) on which rural community uploaded the content. This content included educational material, poem, stories, and advertisement of the local products, local news and tourism. The partners in Nepal used two approaches for content development. One was top-bottom approach and the other was bottom–up approach. In the first approach, the content was produced by ENRD, MPP and other organizations. In the second approach content was produced by the community including teachers, students, villagers and local government. These trainings had been helpful to disseminate research outputs.

3.3.8.5 Infrastructure Development

During both the phases of the project, the country team was provided funding to support monthly stipends for the researchers, acquisition of equipment and books related to different disciplines like linguistics, language processing and computer science. Equipments included PCs, PDA and printers. The availability of funding to support these expenses aided in the development of appropriate localization research infrastructure and enhanced research capacity in the country.

3.3.8.6 Sustainability and Continuity

In terms of sustainability of research, the organizational research capacity was enhanced as the project team graduated from the research on standardization and basic localization during phase 1 to advanced local language computing like script speech and language processing during the project phase 2. In addition, the project trained 31 researchers in the local language computing disciple belonging from different domains like management, CS and linguistics that have later taken the work forward and acquired new projects in the localization research domain.

3.3.9   Pakistan

In Pakistan, the project was executed by CRULP (http://www.crulp.org/) at NUCES, (http://www.nu.edu.pk/) to build research capacity. During Phase 1, CRULP was operating as the regional secretariat (RS) of the project, however during phase 2, in addition to RS; a country project was

also initiated in Pakistan. The following section presents data on each of the six principles of research capacity building impacted through the project.

3.3.9.1 Skill Development

As per the contract, Pakistan country component was required to deliver six localized software namely Localized Email Client, Localized Internet Browser, Localized OpenOffice.Org Writer, Localized OpenOffice.org Writer, Localized OpenOffice.org Draw, Localized Web Composer and Localized Psi. In addition, Pakistan country component also worked on Localized Website Development Tool, Online Stemmer, Machine Translation System, Part of Speech Tagger, Text Normalization Utility, Spell Checker Utility, OpenOffice.org Suite and Psi Chat Tool during the PAN project.

Pakistan country component developed English-Urdu parallel corpus by translating first 100,000 words of PENN Treebank. This work had been supported by PAN Localization project through funding of GSK of Japan. As a part of this project, a Part of Speech tagset was also designed following the PENN Treebank guidelines. Developed tagset contains 46 tags which were properly defined with examples. Finally, translated corpus of PENN Treebank was manually tagged with this tagset.

An additional work had been carried out to computationally enhance Urdu POS tagset for automatic tagging. Tagset was reduced to 32 tags by selection of only those tags which give contextual information instead of lexical information. Urdu corpus translated from PENN Treebank was tagged with this tagset. An additional corpus of 100,000 words from news data is also tagged. The tagging process had been semi automatically carried out with the help of TnT tagger. Trained tagger showed accuracy around 91%. This work was published in 7th Workshop on Asian Language Resources, ACL- IJCNLP in Singapore in 2009 http://crulp.org/research/papers.htm. In 2009, an Urdu rule based stemmer was also developed. In order to cater 140,000 words, 174 prefixes and 712 postfixes were identified. Rules for affix removal are developed with the help of these lists. These rules were helpful for stemming process of Urdu words. The reported accuracy of the system is 91.2 %. Urdu stemmer is also available online http://www.crulp.org/software/langproc/UrduStemmer.htm. Figure 3.37 shows the stemming of one Urdu word.

Figure 3.37: Online Stemming of the Urdu Word "بد دلی"

65

In addition to POS tagger and Stemmer, CRULP had been working on development of Collation sequence, Spell Checking and Text Normalization applications. Some of the work in these areas had already been done and improvements had been partially supported by PAN Localization project. All of these utilities are available online http://www.crulp.org/software/langproc.htm. Pakistan team had also initiated work on development of Urdu OCR system during Phase II. The work started with research and development of multilingual framework for OCR systems. Basic design had been made in this activity. In October 2007, Google released first version of their multilingual OCR framework OCROpus http://code.google.com/p/ocropus/. It was decided to participate in research and development of OCROpus and build working Urdu modules for that system. Work on preprocessing engine had been carried out and a technique for font size independent OCR is proposed.

A working solution for isolated Urdu characters had been made and tested as a part of this project. The process starts by taking the boundary of the image. After outline approximation using splines of the boundary the scaling was applied to have the ligature size according to the 36 font size of the respective ligature. The boundary of outline was filled and then converted into the image form. The reported accuracy of the system is 96 % for single character ligature.
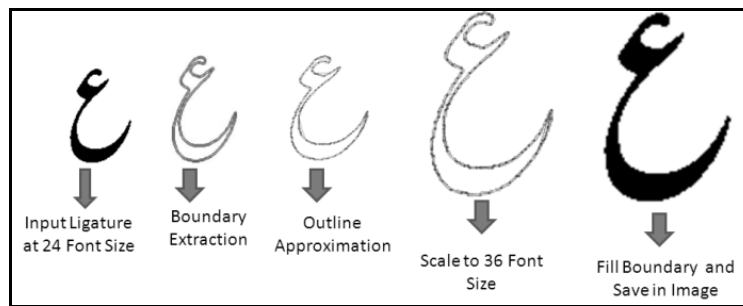
Figure 3.38: Prototype of Urdu OCR

The paper titled "Font size independent OCR for Noori Nasstaleeq" had been presented in Proceedings of Graduate Colloquium on Computer Sciences (GCCS), Department of Computer Science, FAST-NU Lahore, Volume 1, 2010. In addition to that, research on OCR development of Noori Nastaleeq font had been carried out. Prototype of this research was developed using HMM based recognizer. The developed system was an advancement of already working system which had partial alphabet coverage. The system contains different modules of OCR system as segmentation, recognition and post processing units in order to cater all combinations of Urdu characters. The system identifies ligatures of Noori Nastaleeq of 36 point size. The application had been tested for 5,000 frequent words of Urdu and showed accuracy around 80%. This work is partially supported by PAN Localization project. Following figure depicts the interface of application.
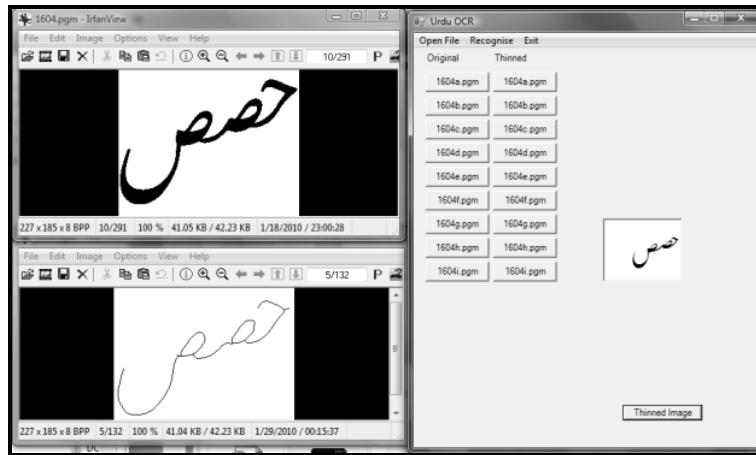
Figure 3.39: Prototype of Urdu OCR

Work had also been done in development of MT system for English-Urdu pair. A rule based MT system has already been developed by CRULP for Urdu. In order to enhance the efficiency and accuracy of that system, research on statistical MT system has been carried out. It is very difficult to cover diversity of Urdu language with rule based system. The proposed work is an enhancement of the system to cater longer and more complex sentences. Research on Semi-Automatic Lexical Functional Grammar Development has been published in the Proceedings of the Conference on Language and Technology 2009 http://www.crulp.org/research/papers.htm.

Figure 3.39: Rule based Machine Translation system

The Urdu Localization Terminology Glossary was developed for the localization of open source software. It is based on the Electronic Dictionary of Localization of Computer Applications (English-Urdu), 2005 by the Center of Excellence for Urdu Informatics, National Language Authority, Islamabad (Pakistan). In

addition to this Mozilla Urdu Language Pack, OpenOffice, FireFox and Thunderbird (for Urdu-India) have also been consulted. This glossary is online available http://www.crulp.org/software/localization/OSS/ossGlossary.html. The figure shows the Urdu translation of the word "Abbreviation" along with the source information from which word translation is taken.
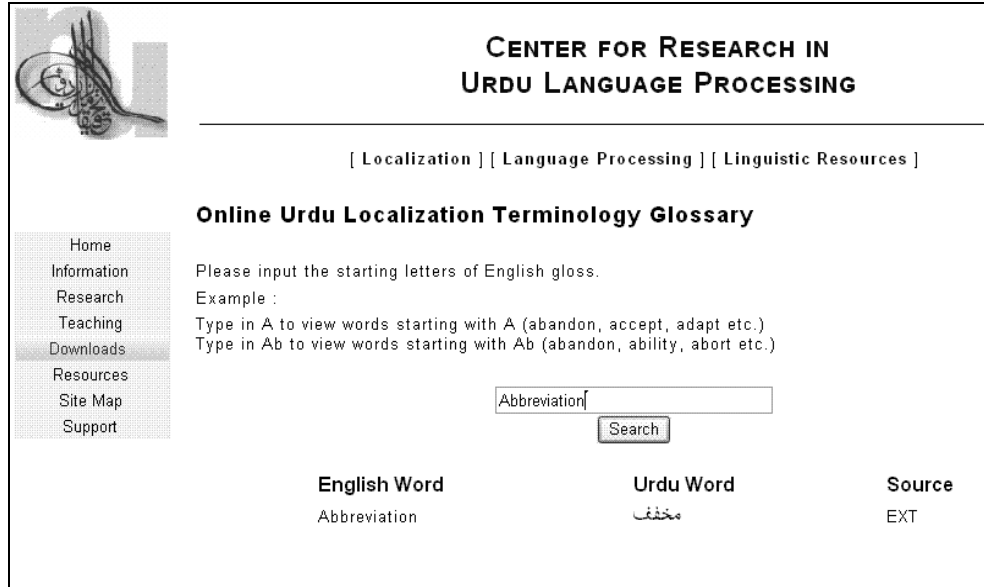
Figure 3.40: Online Urdu Translation of the word "Abbreviation"

The following table summarizes the status of software deliverables required during the project's phase, along with the respective skills required in linguistics, computational linguistics (CL) or computer sciences (CS) to develop the respective software.

Pakistan				
Localized Software	Ling.	CL	CS	Status
Required Localized Software				
Localized Email Client	*		*	Completed
Localized Internet	*		*	Completed
Localized	*		*	Completed
Localized	*		*	Completed
Localized Web	*		*	Completed
Localized Psi	*		*	Completed

Table 3.15: Pakistan Team's Status regarding Localized Software

The table shown above, present that the Pakistan country team was able to develop all the localized software as per the contract. Researchers have focused on the development of basic, intermediate and advanced complexity local language computing applications.

Based on the project experience, the Pakistan country leader was asked to prorate the ability and skill development/enhancement of the organization's researchers during project phase 2 on a scale of 1-5; where 1 represent *Challenged*; 2 represent *Fair*; 3 represent *Average*; while 4 represents *Good* and 5

represents *Excellent* enhancement in the team's performance. The following table presents those figures for assessment of its team's capacity by the project leader from Pakistan country component collected for the team's performance at the beginning of project Phase 2 in 2007 and towards the end of this phase in 2009.

CRULP		
Skill Development	Start of Project, Early 2007	Towards Project End, Mid 2009
LLC Project Development	5	5
LLC Project Design	5	5
Problem Identification	4	4
Project Implementation	4	4
Ability to Do Analysis	4	5
Ability to Communicate Results	2	5
Multi Disciplinary Research	3	5
Quantitative Analytical Skills	3	4
Qualitative Analytical Skills	2	4

Table 3.16 Pakistan Team's Performance regarding Skill Development

The above mentioned table showed that the project leader had confirmed the enhancement of team's skills starting from project development and design to its implementation and analysis within the 3 year span of the project specifically in over-all project execution. According to the project leader, the maximum skill enhancement had been in the project development and in the project design which included the researcher's ability to conceptualize new projects. Teams had also significantly improved in other project areas including the ability to conduct multi-disciplinary research, problem identification and quantitative as well as qualitative analysis of the research.

Ability to publish research in the form of research papers was a salient indicator for measuring the researcher's research capacity. Thus publications of research papers produced by PAN Localization project teams at various national as well as international research conferences was used as the second indicator for analyzing research capacity enhancement. The project team of Pakistan published 5 research papers covering M&T, POS, and IDN during the project. Pakistan was one of those countries who had published maximum number of research papers. Detailed list of research report publication by each country component is presented in Appendix B.

3.3.9.2 Training to Conduct to Practice Research

In Pakistan, CRULP conducted training on localized technology through the Dareecha ICT literacy training program. This training program focused on students from rural areas. The participants were trained on basic computer skills and localized applications, including OpenOffice (Word Processor, Graphics Editor), SeaMonkey (Browser, Email Client, and Web Editor) and Psi (Instant Messaging Client).

In order to establish the need for localized application, specific question was asked from the communities regarding the language that they speak at home and at their work. Answers from this question would

ascertain their preference of language to undertake everyday communication, both written and verbal. When end-users were asked regarding the language spoken at home and work, 100% respondent indicated that they only use local language for communicate at home as well as at their workplace.

This response clearly indicated that the language most convenient for communication for the specific communities was their respective local language. Thus researching for development of local language ICT applications becomes directly useful and relevant to the subject communities, because in order to communicate electronically, and for work, the communities would require applications developed in local languages of the communities.

The respondents were also asked to rate their reading skill and writing skill in English on a scale ranging from Excellent to poor. 61 respondents answered this question and majority of them (20) rated their reading skill in English as excellent and only 2 respondents rated their reading skill in English as poor. A large majority of them (35) rated their writing skill in English as excellent and no one of respondents rated their writing skill in English as poor.

Similarly the respondent were also asked to rate their reading skill and writing skill in Local Language on a scale ranging from Excellent to Poor. 61 respondents in total answered this question and majority of them (24) rated their reading skill in local language as excellent and no one of respondents rated reading skill in local language as poor. 10 respondents rated their writing skill in local language as excellent and nobody rated writing skill in local language as poor. These results are presented in the graphs below.
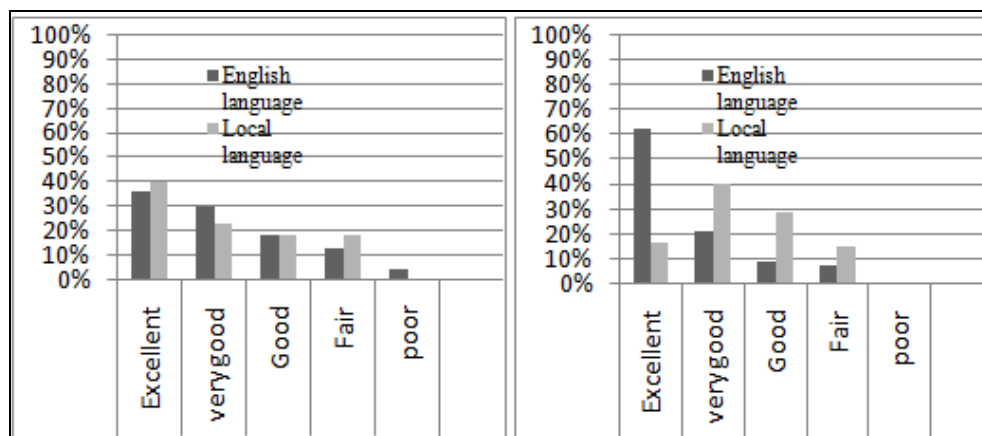
Figure 3.44: Graphs showing reading and writing ability of the respondents in English vs. Local language from Pakistan

These responses from the local communities helped establish the fact that local language computing research initiated by the project is predominantly required by the nation. It is very *close to practice* and urgently required for national development to proliferate the use of ICTs in the country.

3.3.9.3 Development of Linkages

Pakistan country component collaborated with 20 international organizations through the PAN Localization project. These collaborations enabled the partners to collectively plan the technical and financial details, exchange data and technology and discuss and formalize shared intellectual property

regimes, building institutional capacities in the context. The project team has been participating in online research networks.

3.3.9.4 Dissemination

Pakistan country Component has been highlighting its work through its local website (http://www.crulp.org/dareecha/).

Figure 3.41: Website of Dareecha Project

The project had organized awareness seminars to disseminate and publicize research results to local community. These seminars had been attended by a large number of participants from academia, public and private sectors. Through these seminars CRULP had been regularly presenting its work to the key stakeholders from government, IT industry, academia, media, and end user communities. In Pakistan, the country team had been active in interaction with various stakeholders such as Ministry of Information Technology and Telecommunications (MoITT), Ministry of Education, Pakistan Telecommunication Authority (PTA), Universal Service Fund Guarantee Ltd (USF), Pakistan Software Export Board (PSEB), All Software Houses Association (PASHA), Internet Service Providers Association of Pakistan (ISPAK), and Computer Society of Pakistan, National Language Authority NLA (Muqtadra Qaumi Zaban, Urdu Science Board. In these meetings, the country team created awareness about the benefits of localized technology.

CRULP conducted a pilot training of a group consisting of seven 8 grade students in May 2008. The participants were trained on basics computer skills and web browsing. This exercise was beneficial in many ways; it helped the team improve the localized training material, and it also gave the experience to the trainers.

A very large project, named Dareecha, involving a team of 14 human resources has been carried out during Phase II of PAN Localization project. Dareecha project aimed to determine and evaluate sustainable strategies for ICT access and local language content generation in rural schools. The project covered a diverse range of activities including local language resource development, local language

technology deployment and project monitoring and evaluation. Dareecha work can be divided into three phases.

In the first phase, localization of web browsing and emailing (Sea Monkey), word processing (Open Office), chatting (Psi) and web page development (Web Composer) was done in Urdu during the first year using OSSs. It aimed to provide the support for those people who do not understand English but Localization (Urdu translation) can easily make these tools available for them. The training material comprising of seven books was developed in Urdu language and it is freely available online at http://www.crulp.org/dareecha/Training-Books.htm.The internet related applications, SeaMonkey was selected for the localization of the GUI. The reasons behind this tool selection were 1) it is a complete Internet suite including a web browser and an e-mail client (because a browser and e-mail client was integrated into a single application, the overall localization effort would be less as compared to localizing a browser and e-mail client separately), 2) it has Unicode (UTF-8) and bidirectional language support which was required for Urdu, 3) it is localizable, and 4) it was supported on Windows, Linux and Mac OS X. Openoffice.org was selected because of its wide range of usage on multiple platforms including Windows, Linux and Mac. In addition to its universal usage, defined localization process and updated support make the localization process very smooth. Still there were many challenges in this process. Some of the hurdles faced during this project were presence of very large translation strings, selection of appropriate (understandable for layman) meaning and building of office suite. Translation process was carried out with the help of OmegaT translation management tool. In order to extract strings from SeaMonkey suite, Mozilla translator was also used. Different translation resources were consulted to accomplish localization process. These resources included (higher to lower priority) glossary of National Language Authority Pakistan, Firefox 1.0.6 ur-IN glossary and OpenOffice 2.0.3 ur-IN glossary etc.

In the second phase of Dareecha, the material for training of these tools was developed. For the assessment of the training material, competency levels were defined. The competency levels followed a step-by-step approach towards learning new concepts. Starting from beginner's level, difficulty increased with each level. Competency levels enabled us to devise lecture plans for students and later, assessments were also designed on the basis of these competency levels. Each series of competency levels consisted of a set of entry criteria, which must be met before users could proceed with learning, and a set of exit criteria, which must be achieved at the end of the learning associated with that learning area. Within the entry and exit criteria a series of skills was needed to be acquired corresponding to the learning area, and as the users transition through the steps, they were going through increasing levels of difficulty. Each skill level was then further broken down into a more specific skill set. The Competency levels for training of basic computing, web browsing, email, instant messaging, word processing, graphics editing and web page development were defined and made available online at http://www.crulp.org/dareecha/Training.htm.

In the third phase, training for these tools was provided in ten (five boys, five girls) schools of rural areas from 3rd November 2008 and ending on 12 June 2009. The Training sessions for students (grades eight and nine) was conducted across six locations. At least 14 students were nominated by each school to attend the training sessions. Nominated teachers from these schools received trainings along with helping material before the training sessions of students. The purpose was to prepare them for the assistance of Project Dareecha trainers during the student training sessions. There were three five-day training sessions

for students at each school, conducted by Project Dareecha team with the assistance of the trained school teachers. Between the training sessions, the students were left under the supervision of the school teachers. It gave them opportunity to continue practicing the skills they have learnt during training sessions. School teachers were provided with additional training material (work plans, exercises etc.) that can be used between training sessions. Teachers from each school were trained together at a single location whereas students were trained at their respective schools. After the teacher training session, the country team replicated the same training session at each school for the students. After each student training session, the students worked under the teachers' supervision teachers at their school. Training material including books, presentation slides were distributed freely among training participant and books were also provided to each school to use in further trainings.

Figure 3.46: Students during Dareecha Training Sessions

Recognizing the cultural norm, a team composed of exclusively female members conducted trainings in the girls' schools, and a team composed of exclusively male members conducted trainings in the boys' schools. The schools were encouraged to develop business models to use the labs in the evenings to generate income. The program was initially expected to train 140 students but considering the growing demand for training, 228 students were trained. Among them 140 were girls and 88 were boys. 20 teachers were trained. Among them 10 were male teachers and 10 were female teachers. The detail of training program is available at http://www.crulp.org/dareecha/Training.htm.

Figure 3.42: Training Session for Teachers

CRULP team also trained Master trainers/teachers in Punjab IT Labs started by the Punjab Government in 2009 on Urdu usage on computer. In the Punjab IT labs project, IT labs were deployed in 4,286 secondary and higher secondary schools of the province. CRULP contributed to improve the usability of Urdu on computers. A phonetic keyboard layout and fonts (Nafees Nastaliq and Nafees Web Naskh) package was provided for easy use of Urdu on the computer on the MS Server 2008 platform deployed by the Punjab IT Labs Project. The detail of this activity is available at http://www.crulp.org/Downloads/PunjabITLPS/UrduUsageGuide.pdf.

CRULP actively participates in developments taking place in IDNs worldwide. A workshop was organized by CRULP in year 2008, gathering participants representing the various languages spoken in Pakistan http://www.panl10n.net/english/Pakistani_IDNs_Workshop.htm. An initial attempt was made to draft character sets for different languages including Balochi, Pashto, Punjabi, Saraiki, Sindhi and Torwali. A follow-up workshop on IDNs was arranged at NUCES on behalf of the Ministry of IT Pakistan in May 2009, to build on the earlier for Pakistani languages http://www.panl10n.net/english/PakistaniIDNsWorkshop2nd.htm. This workshop was organized in two sessions: the first session was an open discussion where general public was invited through advertisements placed in Urdu and English newspapers. The second session was a closed group meeting where experts finalized decisions on major issues regarding implementation of IDNs for Pakistani languages. CRULP is also a member of the technical committee formed by the Ministry of IT (MoIT), Pakistan to implement IDNs. Plan of a third meeting in collaboration with the MoIT is under way.

In Pakistan, it was strategically planned to develop locally relevant content using bottom - up approach. For this purpose, teachers and students were trained to develop simple web pages. They were encouraged to develop websites. At the end of training, an inter-school web development competition was organized. Entries were invited in three categories: Community website (by a group of students), School website (by a group of teachers) and Individual website (locally relevant website by individual students). 57 entries in total were received, 10 websites were developed by teachers and 47 were developed by students. Women also participated in significant number in the competition. Among 47 websites developed by students, 36 were developed by female students whereas 11 websites were developed by male students. All websites are available at http://www.crulp.org/dareecha/competitions/competitions.htm.

 These websites were evaluated by a panel of judges from the IT industry, academia, government, media and other organizations. After evaluation of websites, a prize distribution ceremony was held on January 23, 2010 at the NUCES Lahore campus and 250 project participants attended the ceremony. The websites awarded with prizes are available at http://www.crulp.org/dareecha/winners.htm.

Figure 3.43: Homepage of Website winning First Prize

The Pan Localization Pakistan component Dareecha had been presented at the following events at national and international levels to disseminate their research outputs:

Gendered OM methodology selected for demonstration at International Conference on Information and Communication Technologies and development ICTD2009, in CMU Doha, Qatar from April 17-19, 2009.

Sustainable Development Policy Institute (SDPI) Study Group Meeting on Women and ICTS: Exclusion or Empowerment on 13 August, 2009. Internet Governance Forum (IGF) Workshop on Equality in access to knowledge society, through language and cultural diversity held at Sharm El Sheikh, Egypt on 18 November, 2009. 2nd GEM Global Exchange, will be held in Bali, Indonesia from November 24-30, 2009.

Seminar on Integrating IT in Education: Language, Curriculum and Training Challenges in Government Schools of Punjab held on 17 December, 2009 at NUCES, FAST Lahore.

In addition, the research paper on Urdu stemmer had been presented by Dr. Sarmad Hussain at 7th Workshop on Asian Language Resources, ACL- IJCNLP in Singapore in 2009. David M. Malone, President of Canada's International Development Research Centre, IDRC president visited CRULP on 5 February, 2010 and highly appreciated the achievements of country team. The project activities in Pakistan have been widely covered by print and electronic media. Spider, popular IT Magazine in Pakistan published articles highlighting progress made by the project in the country.

3.3.9.5 Infrastructure Development

During both the phases of the project, the country team was provided funding to support monthly stipends for the researchers, acquisition of equipment and books related to different disciplines like linguistics, language processing and computer science. Equipments included computer hardware like PCs, laptop, scanners, printers, CD-RW Drive, USB Drive and mobile phone for end user. Pakistan country component also focused on development of networking and available funds were mostly used for these

activities. The availability of funding to support these expenses aided in the development of appropriate localization research infrastructure and enhanced research capacity in the country.

3.3.9.6 Sustainability and Continuity

In terms of sustainability of research, the organizational research capacity was significantly enhanced as the project team researched in standardization, basic localization, Language processing, script processing and speech processing during PAN Localization project. In addition, the project trained 44 researchers in the local language computing disciple belonging from different domains like management, CS and linguistics that have later taken the work forward and acquired new projects in the localization research domain.

3.3.10   Sri Lanka

In Sri Lanka, the project was executed by the UCSC, http://www.ucsc.cmb.ac.lk/.

The usage of computer per 100 populations in Sri Lanka was 8.2 in 2008 (Weerasinghe, R & Desilva, C, 2009). The figure depicts that the usage of computer in Sri Lanka was comparatively more than other countries discussed above. However, development of localized software could be helpful to maximize the use of ICT because the majority of the people speak in their local language "Sinhala". According to Statistics (Scenic Sri Lanka, 2007), "Sinhalese make up about 74 percent of the Sri Lankan people and the language they speak is Sinhala, which is an official language". When technology is available in their local language then much more people would easily use latest technology. Through PAN Localization project, Sri Lanka country component worked on the development of localized software and this project boosted them to build research capacity in local language computing. The following section presents data on each of the six principles of research capacity building impacted through the project.

3.3.10.1 Skill Development

During Phase I of the project, Sri Lanka country component was required to develop Encoding Conversion Utility, Sinhala Lexicon, Multilingual Lexicon (Sihala, English, Tamil), Sinhala TTS System and Sinhala OCR System. In addition to the agreed software, Sri Lanka country component also developed Sinhala Unicode Converter, Sinhala Syllabification Tool, Sinhala OCR Preprocessor, Sinhala Optical Character Recognition System and Sinhala Text to Speech System.

To enhance skill development of Sri Lankan team, training on Phonetics and Phonology for TTS, was conducted at the UCSC, Colombo, Sri Lanka, by Dr. Sarmad Hussain, from 21st - 25th June 2004. Objectives of the training were to give the requisite background to the Sri Lankan PAN Localization team and others interested on Phonetics, Phonology, Acoustic Phonetics and how all this fits in with the TTS process.

Figure 3.44: Training on Phonetics, Sri Lanka ,2004

Sinhala Text to Speech system was awarded as the "Most Innovative Product" at the Biennial Infotel Trade Exhibition held in Colombo, Sri Lanka on 1st November 2008 http://www.itpro.lk/node/1554.

During the phase II of the project, List of gTLDs and ccTLDs in Sinhala, English- Sinhala Parallel and Aligned Tagged Corpus 100k words, Sinhala Wordnet 5000 words, TM Application, Sinhala H/W Recognition System for PDAs and Tamil Language Learning Tool were required localized software to deliver. The project team all worked on translation of PENN Treebank to build a Parallel Corpus of Sinhala and English. First 100,000 words had been translated and released as an output of Parallel Corpus. Issues faced during the translation process had also been documented. The work on Part of Speech Tagset and Tagger had already been carried out by Language Technology Research Laboratory (LTRL) and a paper titled "A Stochastic Part of Speech Tagger for Sinhala" had been published in Proceedings of 6th International Information Technology Conference. Colombo, Sri Lanka (2004). A complete document containing tags, meanings and examples had been released as a part of parallel corpus activity. Sinhala version of parallel corpus had also been tagged using developed stochastic part of speech tagger. In addition to that, 100,000 words from 10 Million words corpus of Sinhala had been annotated with POS, resulting a tagged corpus of 200,000 words in total.

The team had also worked for the development of the Sinhala WordNet. For this, different Sinhala sources such as Maha Sinhala Sabdakoshaya by Dr. Harishchandra Wijetunge (Main Source), Sinhala-English Dictionary by Charles Carter, Sinhala-English Dictionary by Benjamin Clough, Sinhala-English Dictionary by Dr A. P. De Soysa, Sanskrit- English Dictionary by Monier Williams, Technical Glossaries published by the Official Language Department, Sinhala Namawaliya (An ancient Sinhala thesaurus) and Ruwanmala (An ancient Sinhala thesaurus) have been used. Previously, very little effort was made for the semantic aspect of Sinhala language. Therefore some English sources were used to extract the semantic aspect of words and their relations and classifications. These sources are Princeton English WordNet (Fellbaum, 1998), Roget's Thesaurus (Roget, 1962) and Suggested Upper Merged Ontology (SUMO) Ontology www.ontologyportal.org.

1100 high frequency synsets (Approximately 5,000 words) had been selected from the UCSC Sinhala corpus. All those words had been considered as the separate entries for the WordNet which had different semantic information. In the next step, sense identification process had been performed in which a linguist had determined the list of senses from the dictionary and English WordNet. Maha Sinahla

Sabdakoshaya and the Princeton English WordNet were the main sources for this purpose. All the sense relations had been extracted from the database instead of defining them from scratch. This procedure helps to develop the Sinhala WordNet using relations without duplication of effort. After this, these relations had been stored in a human readable form. The Sinhala word for each entry had been translated into the English, and then the synset identifier of the translated word obtained from Princeton English WordNet. The Sinhala word along with the synset identifier is then inserted into the Sinhala WordNet. In addition to wordnet, a trilingual dictionary (Akaradi) had also been made for English, Tamil and Sinhala users in phase I. These resources are available at http://www.ucsc.cmb.ac.lk/ltrl/?page=services&lang=en&style=default.

Sri Lanka team had worked on language table and terminology translation of gTLDs and ccTLDs in Sinhala for IDNs. The translated Sinhala gTLDs and ccTLDs can be accessed online http://www.panl10n.net/english/OutputsSri-Lanka2.htm. These efforts had contributed in posting a formal application to ICANN for registration of Sinhala and Tamil country codes. It had been drawn from work carried out by ICTA with technical support from University of Colombo. ICTA of Sri Lanka and the LK Domain Registrar had been part of the global discussion on the fast-track IDNs process. An ICTA commissioned working group, consisting of experts from within itself, the UCSC, the UOM and the LK Domain Registry (a subset of its Local Language Working Group) has been looking into issues relating to this process. Major issues had also been discussed such as homoglyps, mixing of scripts, multiple forms of the same word, spelling variants and browsers issues. In addition, a systematic process, considering both human and technical factors, for providing a sustainable IDNs solution for languages belonging to the Indic family has been developed. Sri Lanka developed Sinhala and Tamil IDNs in June, 2010.

Sri Lanka team had also been working on developing web applications in order to disseminate the research and obtaining feedback from common users. A useful application of this category is EnSiTip which is a Firefox Plug-in http://www.ucsc.cmb.ac.lk/ltrl/projects/EnSiTip/. It facilitates users by providing Sinhala translations for English words in a popup menu for any web page being browsed. The plug-in has been downloaded 15,942 times until 3rd March, 2010. In addition to this plug-in Online Encoding Converters for Sinhala had also been made available for public usage http://www.ucsc.cmb.ac.lk/ltrl/services/feconverter/. Sinhala typing had been made easy by providing an online version of a Unicode keyboard. Text can be written online using this keyboard and copied to any desired text box. This keyboard can be accessed at http://www.ucsc.cmb.ac.lk/ltrl/services/keyboard. A Sinhala Spell Checker (Subasa) was also developed through the capacity built up during the PANl10n project.

Sri Lanka team had also been working on development of Translation Memory tool. The purpose of this application was to facilitate translation process from English to Sinhala with the aid of suggestions and central memory. The open source application OmegaT was chosen for this purpose because of source code availability but OmegaT does not support Sinhala Unicode Rendering (as it has been developed in Java Swing). Therefore, OmegaT's core modules were used to create a web application which supports most of the scripts.

A central system OpenTM, based on OmegaT, had been made in Language Technology Research Laboratory to facilitate the translation process. The system is able to maintain the centralized memory

backup and translations are updated after feedback from translators. Following figure shows the working of this system.
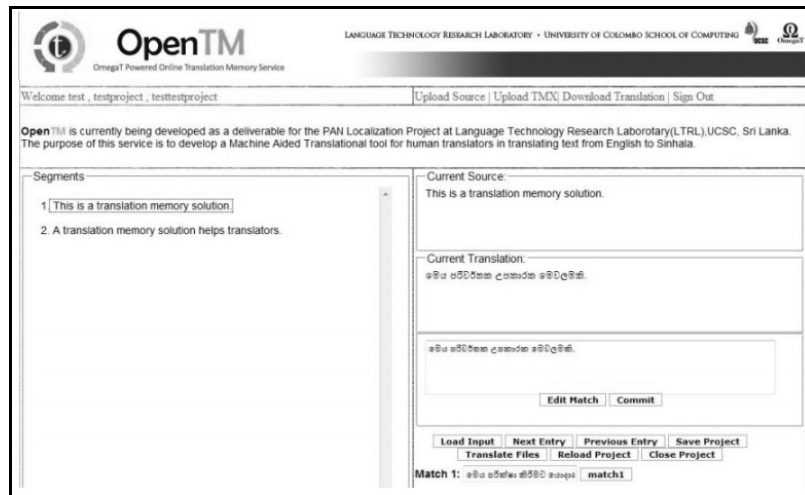
Figure 3.45: Open TM (An OmegaT Powered System)

Another project developed during the phase II was Tamil Language Learning Tool. It had been a very extensive exercise which included development of training content for Tamil in Sinhala and English (Official languages of Sri Lanka). The content was generated with the help of an already existing book like "An Introduction to Spoken Tamil by James W. Gair, S. Suseendirarajah and W. S. Karunatilaka" which has been translated in Tamil by Rev. Kadurugamuwe Nagitha Thero, a lecturer at the University of Kelaniya, Sri Lanka. Interactive questions are selected to cover daily life scenarios in the form of 25 lessons. There are three different sections of each lesson; Dialogs, Grammar and Exercise (For details please see Sri Lanka section of Content chapter).

A parallel activity of application development had also been carried out along with learning content generation. Complete training data is stored in XML format which is used by the Language Learning Tool. For development purpose, Macromedia Flash had been selected because of its support of Unicode and portability on desktop and websites. The tool facilitates learner with Sinhala and English interfaces for all lessons. Following figure displays the application layout.
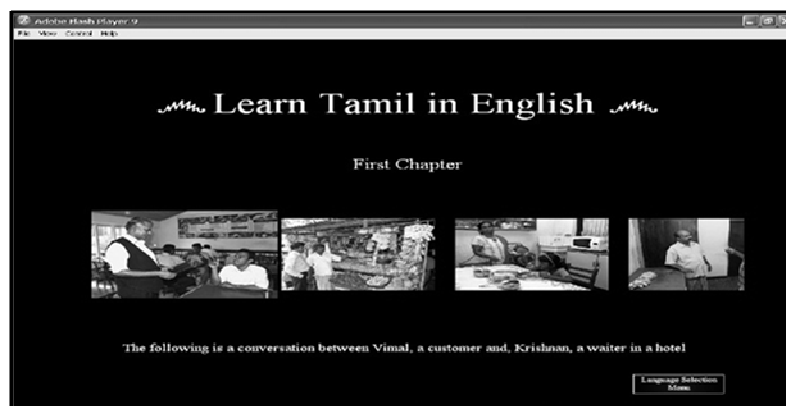
Figure 3.46: Tamil Language Learning Tool Application

In addition to this tool an online tool for building a trilingual glossary (English-Sinhala-Tamil) had also been developed which provided the facility to vote for each term to be translated.

Sri Lanka team had also initiated research on Online Handwriting Recognition system. The objective of the work was to develop an isolated recognizer for real time input. The proposed model was based on Template Matching with Dynamic Time Warping (DTW) distances. The project reached its first milestone and data capturing, data analysis and classification modules had been developed. Training data had also been collected from 12 different users on PDA. The training set contains three samples of 1,500 strokes. An error analysis had been carried out on collected samples to prune the training data. The research on data cleaning and system training was being continued after the completion of PAN localization project.

Sinhala Text to Speech system was developed during Phase I. In second phase a screen reader facility had been incorporated in this TTS application. All of the outputs discussed and relevant research material is available on PAN Laicization website http://panl10n.net/english/OutputsSri-Lanka2.htm and regular updates can be found on http://www.ucsc.cmb.ac.lk/ltrl/.

The following table summarizes the status of software deliverables required during the project's Phase 1 and 2, along with the respective skills required to develop the respective software.

Sri Lanka				
Localized Software	Ling	CL	CS	Status
Required Localized Software of Phase I				
Encoding Conversion Utility			*	Completed
Sinhala Lexicon		*	*	Completed
Multilingual Lexicon		*	*	Completed
Sinhala TTS System		*	*	Completed
Sinhala OCR System		*	*	Completed
Other Localized Software of Phase I				
Sinhala Unicode Converter				
Sinhala Syllabification Tool				
Sinhala OCR Preprocessor				
Sinhala Optical Character				
Sinhala Text to Speech System				

Table 3.17: Sri Lanka Team's Status regarding Localized Software of Phase I

Required Localized Software of Phase II				
Localized Software	Ling.	CL	CS	Status
List of gTLDs and ccTLDs in	*		*	Completed
English- Sinhala Parallel and Aligned Tagged Corpus 100k	*		*	Completed
Sinhala Wordnet 5000 Words	*	*	*	Completed
TM Application	*	*	*	Completed
Sinhala H/W Recognition	*		*	Completed

Tamil Language Learning Tool	*	*	*	Completed
Other Localized Software of Phase II				
Linux Windows				
Tamil Language Learning tool				

Table 3.18: Sri Lanka team's status regarding Localized Software of phase II

The above mentioned table showed that the project team had been able to submit all localized software as per the contract and the 100 % accomplishment of the localized software showed that Sri Lanka country team's skill had enormously enhanced over the project implementation.

Based on the project experience, the country project leader was asked to prorate the ability and skill development/enhancement of the organization's researchers during project phase 2 on a scale of 1-5; where 1 represent *Challenged*; 2 represent *Fair*; 3 represent *Average*; while 4 represents *Good* and 5 represents *Excellent* enhancement in the team's performance. The following table presents those figures for assessment of its team's capacity by the project leader from country component collected for the team's performance at the beginning of project Phase 2 in 2007 and towards the end of this phase in 2009.

Skill Development	Year 2007	Year 2009
Project Development	2	4
Project Design	2	4
Problem Identification	3	4
Project Implementation	3	4
Analysis	3	4
Quantitative Analysis	2	4
Qualitative Analysis	3	4
Multi Disciplinary Research	2	5
Communication	3	5

Table 3.19: Sri Lanka team's performance regarding skill development

The publications of research papers produced by PAN Localization project teams at various national as well as international research conferences is used as the second indicator for analyzing research capacity enhancement. The project team of Sri Lanka published 10 papers covering MT, Lexicon, Speech and IDN during the project's phase 2 which is the highest among the all participating countries. Detailed list of research report publication by Sri Lanka project team is presented in Appendix A.

3.3.10.2 Development of Linkages

Sri Lanka country component collaborated with six organizations at nation level namely Information and Communication Technology Agency, Parliament of Sri Lanka, University of Peradeniya, Sabaragamuwa University, University of Kelaniya, University of Colombo and Sri Lanka Collage of Journalism.

3.3.10.3 Dissemination

The Sri Lanka county component disseminated its work through its website http://ucsc.cmb.ac.lk/ltrl/.
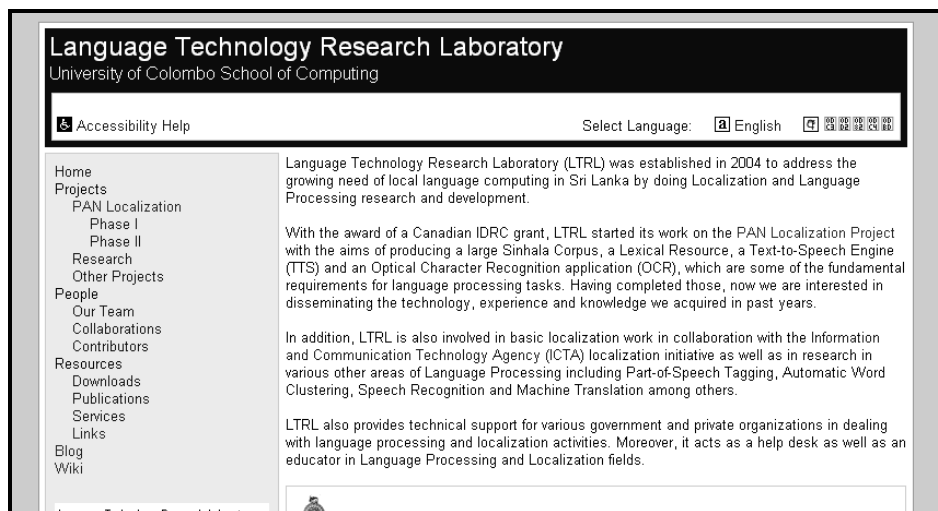
Figure 3.47: Website of LTRL

The country team had been at the centre of efforts to develop local content. Sri Lanka component conducted several trainings for University staff and fresh graduates to disseminate their research outputs. The training participants were trained to write articles to Sinhala Wikipedia and blogs. The main objective was to the increase the local language contents on the web. University staff was from the University of Sabaragamuwa while students were following computer awareness program under IRQUE project http://www.irque.lk.

Trainings for university staff was held in two stages. One workshop was conducted for academic staff and at the other was conducted for non-academic staff. These trainings covered the topics including Sinhala Unicode, Structure of Internet and Web 2.0, Wikis and blogs. The country component developed a LLT in Sinhala and English to facilitate Tamil language learning. The main objective of this project was to make language learning process less arduous and the framework followed is flexible enough to extend it to be used by other project partners to teach their languages. The tool contains twenty five lessons. Each lesson has dialog, grammar section and exercises. First five chapters describe the basics in the Tamil language. Exercises can be found at the end of each chapter and reviews can be found after every five chapters. The content for this tool was taken from the Sinhala translation of "An Introduction to Spoken Tamil" by James W. Gair, S. Suseendirarajah and W. S. Karunatilaka. The translation work was done by Rev. Kadurugamuwe Nagitha Thero, a lecturer at the University of Kelaniya, Sri Lanka. According to the author, the book increases awareness about Tamil language and also gives an idea about its culture. The book is helpful to learn spoken as well as written Tamil. Tamil LLT material is available at http://www.panl10n.net/english/OutputsSri-Lanka2.htm.

Figure 3.48: Main Page of Language Learning Tool

The country team conducted its first training on Introduction to Unicode and Sinhala Wikipedia for 55 undergraduates of University of Colombo on 26 March, 2008. Second training was conducted for 51 undergraduates on 10 June, 2008. Third training on Introduction to Web 2 and its application in local languages was conducted for Academic staff, University of Sabaragamuwa, Sri Lanka on 12 June 2008. 30 participants were trained in this training. Fourth training on Introduction to Web 2 and its application in local languages was conducted for Non-Academic staff, University of Sabaragamuwa, Sri Lanka on 13 June 2008. 48 participants were trained in this training. Fifth training on Introduction to Unicode and Sinhala Wikipedia was conducted for 24 undergraduates on 22 July, 2008. Sixth training on Introduction to Unicode and Sinhala Wikipedia was conducted for 25 undergraduates on 6 March, 2009. Some of the details are available at http://ucsc.cmb.ac.lk/ltrl/blog/.

Figure 3.49: Non Academic Staff Training at University of Sabaragamuwa on 13 June, 2008

The country team was also part of DAISY related trainings from the beginning, and focused on DAISY compatible digital book production in local languages. Training was conducted at University of Peradeniya, at the request of Professor Weerakkody.

The Sri Lankan team also won "Most Innovative Product" award for Sinhala Text-to-Speech System at the Biennial Infotel Trade Exhibition held in Colombo, Sri Lanka on 1 November, 2008 http://ucsc.cmb.ac.lk/ltrl/blog/. Project partners had been involved in designing, developing and disseminating the material developed, which has contributed to mutual capacity to disseminate research.

3.3.10.4 Infrastructure Development

During both the phases of the project, the country team was provided funding to support monthly stipends for the researchers, acquisition of equipment and books related to different disciplines like linguistics, language processing and computer science. Equipments included PCs. In the operational field, participants regarding different domains of the PAN Localization project were trained. Funds also utilize on software namely Visual Studio.Net and Win Snoori. The availability of funding to support these expenses aided in the development of appropriate localization research infrastructure and enhanced research capacity in the country.

3.3.10.5 Sustainability and Continuity

In terms of sustainability of research, the organizational research capacity was enhanced as the project team researched on standardization, Language processing and Script processing during PAN Localization project. In addition, the project trained 14 researchers in the local language computing disciple belonging from different domains like management, CS and linguistics that have later taken the work forward and acquired new projects in the localization research domain.

3.4 Synthesizing Research Capacity Building in Partner Countries

The following section summarizes the finding presented above and provides a regional perspective on the impact of research capacity building through PAN Localization project.

In terms of Skill development and enhancement, the extent of completion of agreed deliverables, as per the contract, has been used as a measure for illustrating RCB through the project. These results are graphically represented in the figure below.
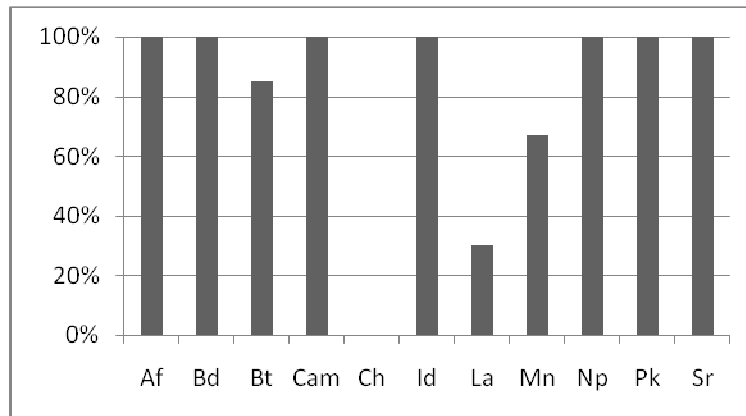
Figure 3.50: Status of completion of deliverables as per the country contracts

Findings show that seven out of eleven country teams were able to achieve the required targets of localized software development as agreed initially in the contracts. Four countries that could not achieve 100% target were Bhutan, Mongolia, China and Laos.

Bhutan and Mongolia have not been off the target in practice. This is because during the course of project, their project plan changed from that committed initially therefore instead of completing the software decided at planning phase the project either discontinued the development of certain deliverables in consultation with the regional secretariat, e.g. the development of speech corpus in Bhutan or replaced the development of a certain software with another software development, e.g. instead of localizing OO, spread sheet and presentation, Sea Monkey suite has been localized in Mongolian. However, a real lag has been observed in the accomplishment of specific research targets in Laos and China. This was because of significant communication and coordination problems with the research team in China and large turnover rate in Laos which resulted in a completely new and inexperienced, even after persistent technical training was being provided to the country component.

To illustrate skill enhancement to conduct local language computing research in partner countries, the other measure used to determine the extent of RCB through the project is the number of research papers published by the country teams. The following table presents the total number of papers published by the project teams during the project's phase 2.

Country Team	No. of Papers	Focus of the Publication
Bangladesh	8	Machine Translation, Script and Speech Processing
Bhutan	1	Text to Speech Synthesis
Indonesia	2	Statistical Language Processing, Machine Translation, Parts
Mongolia	6	POS, Corpus, Speech
Nepal	1	NLP
Pakistan	5	IDN, POS, M&E
Sri Lanka	10	MT, Lexicon, Speech, IDN

Table 3.20: Papers Published

Research teams in Sri Lanka, Bangladesh, Pakistan and Mongolia have produced the maximum number of research papers and on advanced topics of localization research. A closer observation of their team structure reveals two common factors that may be deemed detrimental for such a significant research focus. Firstly, all the four teams have been housed within universities and secondly the educational qualification of their respective country team leaders is doctorate. The educational qualification of the country project team leader has been noted as a significant factor influencing a stronger research focus and thus RCB within the respective teams.

The skill enhancement of the project teams to conduct local language computing research can also be analyzed through the comparison of research targets planned in phase 1 and phase 2 respectively.

During the project planning phases (both phase 1 and phase 2) the country teams were provided freedom to set their own research goals in different domains of the local language computing namely standardization, basic Localizations, language, script and speech processing, .in terms of the software they

chose to develop through the project in consideration of their local capacity and national priorities. The planned research targets during phase 1 and phase 2 are presented in the table below.

Country Team	Standardizatio	Basic	Lang.	Script	Speech
ACSA,	√	√	X	X	X
CRBLP,	√	√	√	√	√
DIT, Bhutan	√	√	X	√	X
MoEYS, Cambodia	√	√	√	√	√
BPPT, Indonesia	X	√	√	X	X
UI, Indonesia	X	√	√	X	X
NAST, Laos	X	√	X	√	X
MUST, Mongolia	X	X	√	X	√
NUM, Mongolia	√	√	√	√	√
InfoCon, Mongolia	√	√	X	X	X
MPP, Nepal	X	√	√	√	√
CRULP, Pakistan	√	√	√	√	√
UCSC, Sri Lanka	√	X	√	√	X

Table 3.21: Country Wise Status on Various Technologies

A longitudinal comparison of the localized software developed by the country teams during project's phase 1 and phase 2 presented in table below show another interesting result.

Country	Basic		Intermediate		Advanced	
	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2
Afghanistan	*	*		*		
Bangladesh	*		*	*	*	*
Bhutan	*	*	*	*		*
China		*		*		*
Cambodia	*	*	*	*		*
Indonesia						*
Laos	*		*	*		*
Mongolia		*		*		*
Nepal			*			*
Pakistan		*		*		*
Sri Lanka	*		*	*	*	*

Table 3.22: Country wise research targets planned in phase and phase to categorized on the basis of complexity levels

Data regarding the research targets accomplished based on the complexity level of localized application development show that the country teams have significantly gained more technical skills over the project implementation. This is evident from the fact that during phase 1 only two of eleven countries were doing advanced local language computing while during phase II 10 out of 11 countries successfully developed advanced local language software.

Research Domain	Year 2007	Year 2009
Project Development	2	5

Project Design	3	5
Problem Identification	3	5
Project Implementation	3	5
Analysis	3	4
Quantitative Analysis	2	4
Qualitative Analysis	2	4
Multi Disciplinary Research	2	4
Communication	2	4

Table 3.23: Comparative analysis of the research skills of the project teams

The same has also been re-affirmed by the country project leaders. When the project leaders were asked to assess their teams' capacity at the beginning of project's Phase 2 in 2007 and towards the end of the phase in 2009, he average results of their responses, shown in the table 3.23 above, present an increase in their research skills. The data presented in the table above provides the rating of the country project leaders in terms of the aforementioned research skills on a 1-5 point scale where 1 represents challenging and 5 represents excellent.

In respect of networking and development of partnerships and linkages, table 3.24 presents the reported number of formal organizational collaborations made by the country teams both at national as well as international level.

Country Team	Number of Collaborations		Total
	International	National	
ACSA, Afghanistan	1	1	2
CRBLP, Bangladesh	-	1	1
DIT, Bhutan	1	1	2
MoEYS, Cambodia	-	2	2
BPPT, Indonesia	-	1	1
UI, Indonesia	1	1	2
MUST, Mongolia	-	3	3
NUM, Mongolia	-	2	2
InfoCon, Mongolia	-	3	3
MPP, Nepal	2	-	2
CRULP, Pakistan	20	-	20
UCSC, Sri Lanka	-	6	6

Table 3.24: No. of organizational collaboration by the country teams

Partner teams were encouraged to establish collaboration with institutions that had more expertise in a specific field. These collaborations enabled the partners to collectively plan the technical and financial details, exchange data and technology and discuss and formalize shared intellectual property regimes, building institutional capacities in the context. It is encouraging to note that every team had initiated collaboration with relevant national and international organizations.

In terms of enhancing the research dissemination capacity of the country teams, the following table presents the number of seminars organized by the country components to disseminate the research work.

Country Team	No. of Events
Afghanistan	1
Bangladesh	2
Bhutan	2
Cambodia	1
Indonesia	1
Laos	1
Mongolia	1
Nepal	1
Pakistan	4

Table 3.25: Country Wise Status of Events/Seminars organized by the teams

These seminars have been attended by a large number of participants from academia, public and private sectors. Through these seminars partner institutions had been regularly presenting their work to the key stakeholders from government, IT industry, academia, media, and end user communities. Such seminars were conducted during both phases of the project. Pakistan organized maximum seminar to disseminate the research work followed by Bangladesh and Bhutan.

Regarding sustainability and continuity of the research capacity, through the PAN Localization project significant numbers of technical developers, linguists and social scientist have been trained. The following table provides the numbers of researchers engaged and trained through the project.

Country Team	Mgmt	Tech	Lang	Social Sc.	Total
Afghanistan	2	2	3	0	7
Bangladesh	3	10	3	6	22
Bhutan	1	8	4	0	13
Cambodia	7	39	13	3	62
China	4	11	3	0	18
Indonesia	3	3	12	0	18
Laos	5	9	6	3	23
Mongolia	3	15	9	0	27
Nepal	6	12	4	9	31
Pakistan	6	27	7	4	44
Sri Lanka	3	7	3	1	14
	43	143	67	26	279

Table 3.26: No. of Researchers trained through PAN Localization Project

Cambodia, Pakistan and Nepal produced maximum number of researchers in the field of local language computing. Also the maximum numbers of researchers engaged in the project belong to a technical/CS background, which is true and required for the accomplishment of the research targets and to develop a sound bass for promoting this research in the region. However, data further shows that there has been a minor focus on the inclusion of social scientist in the project teams, due to which the minimum of social scientist have been trained through the project. Country specific results show that teams that had a larger number of researchers from the social sciences domain within their team have done a more grass root

centric operation, by deploying the technologies developed to the end-user communities, than those who did not include more researchers from social sciences domain.

PAN Localization project has significantly contributed in developing the research capacity in local language computing in each country team. Each team had started to develop the technology in local languages. This transformation would contribute a lot in effective utilization of technology in national development.

3.5   Evaluation of Research Capacity Building Strategies

PAN Localization project aimed to build capacity for research in localization in each of its partner countries. Prior to initiating the capacity development program, baseline study of the existing research capacity in partner countries was conducted to help in devising relevant strategies (Hussain, 2004). The study showed that while teams had very limited experience in basic localization (except in a couple of countries) and many countries had no work done in localization policy development and development of intermediate and advanced localization. Also there was hardly any experience in inter-disciplinary research and development across computing, engineering and linguistics. Regarding team management, only two country team leaders had experience in running long-term multi-disciplinary projects.

Faced with the above capacity challenges, appropriate measures had to be undertaken, focusing across the six principles of research capacity building to target holistic improvement. PAN Localization project's regional secretariat had a focused approach on capacity building the teams in ensuring sustainability of the research being conducted. In this context, the project had instituted specific strategies that are detailed below:

In terms of skill enhancement, the country teams tremendously worked hard to realize their agreed research goals, however the regional secretariat has also steered the overall effort by instituting relevant measures when they were challenged and stopped short of skills to meet the planned research outputs. In this context, four specific strategies were used to facilitate teams in filling the knowledge gap to achieve their deliverables. These are explained in the section below.

Training through Summer School in Local Language Computing

RS had envisioned that the project plans and research goals in PAN Localization project phase 1 were focusing more on basic or intermediate complexity localization software applications. However in project's phase 2, the teams were developing more advanced localized applications for which teams would require advanced skills in local language computing and sound theoretical background.

Thus to facilitate the teams in such development, an innovative form of their technical training called Summer School in local language computing was organized during the time when the country teams were transiting from completion of the phase 1 to the inception of the project phase 2. This was a semester equivalent (three month long) extensive academic program with credit for five graduate courses in linguistics and computational linguistics that were not offered in the partner countries. This program was only offered to those team members who agreed to work on the project for at least one year after the completion of the training.   In addition, as an incentive to go through the training, it was arranged in

collaboration with other universities in South Asia that the credit hours earned through the semester were transferrable in any other graduate program in those universities. The course instructors selected to teach these courses were experts in their fields chosen from around the world. This helped quickly boost the capacity of the partner teams, enabling the transition from undertaking research in localization in Phase 1 to more advanced research in language computing in Phase 2.

3.5.1.1 Short Term Training

Short term training was another capacity building strategy designed by RS. Short term training were organized as a week-long activity targeting training on a specific research area in which the team was lacking in technical competence. Six short term training were conducted during the project, covering a varied set of topics, for example, FOSS localization, OCR development, linguistics and monitoring and evaluation using outcome mapping framework.

As a specific example a five day long training of the Afghanistan country team was held in Pakistan. http://panl10n.net/english/Afghanistantraining.htm. During this training, the team was trained on outcome mapping framework for planning, monitoring and evaluating their projects. In addition specific sessions were conducted on font development and open source software localization that helped the team to initiate their country project.

In addition to building individual's capacity, this mode of training also helped build institutional capacity. Trainees receiving the short term training were not limited to project staff only but would also include additional relevant staff where this training was organized.

3.5.1.2 Mentor Placement Programs I & II

Where the country team required longer training to address capacity challenges, mentor placement programs were initiated through the RS. Through this program a mentor would be place with the partner country that would provide technical and management support to the recipient team. Two different models were adopted in this context. In first model (referred to as mentor Placement I in Table 3.27), a mentor from within the partner countries was sent to partner needing support. Three such mentor placements were conducted from 2004-2007, and 2 were held during the second phase of the project.

With the mentor placement held in Bhutan during the project phase 1, country team was guided by Mr. Guntupalli Karunakar, on localization of Linux and Open Office in Dzongkha. Following this training, the country team was able to develop a live debian based Dzongkha Linux distribution.

In second model (referred to as mentor Placement II in Table 3.27) respective country component nominated one or two persons from team to stay with mentoring organization for the training duration. One such placement was initiated in the project's first phase of the project, while 5 such placements were done in the Phase 2. While both models worked out equally well, an extension of first model has also been tried by providing the remote mentoring facility after completion of training, which has also proved effective in achieving the research outcomes.

The following table summarizes the number of times each intervention was conducted during the project's phase 1 and phase 2.

Training Strategy	Phase 1 (2004-07)	Phase 2 (2007-10)
Short Term Training	6	-
Mentor Placement (I)	3	2
Mentor Placement (II)	1	5
Summer School	1	-
Regional Training	-	2

Table 3.27: Capacity Building Interventions during Project Phases 1 and 2

As presented in the table above, it is evident that the project's Phase 1 focused more on short term training and mentor placements I. However as the teams gained more in their technical skills the project's Phase 2 strategies targeted collaborations, e.g., Mentor Placement II, summer school and conference participation, and longer term impact, e.g. through support for higher studies.

3.5.1.3 Cost, Duration and Capacity Gain Thresh-holds

A quantitative view of deploying each strategy can be computed based on the (approximate) average cost incurred for each instance of the strategy employed, its duration and the intended maximum capacity gain that the strategy was able to elicit. Based on the experience of the PAN Localization project *phase 1*, the following numbers have been computed in the table below that provides a starting point for cross-comparison of strategies on salient factors.

Strategy	Cost (In 1000 of USD Per Instance)	Time (Upper Limit)	Unit Cost (Per Person, Per Month, Per Instance)	Potential Capacity	Interpretations
Deliverable Development	5.1-6.8	36 months	179	4	Able to Implement the Solution
National Training	2	1 week	1143	2	Able to Understand the Problem
Mentor Placement	7.6	6 months	217	4	Able to Implement the Solution
Regional Training	100	1 week	13333	1	Able to Understand the Problem
Summer School	133	3months	1478	3	Able to Understand the Solution

Table 3.28: Cost, duration and capacity gain thresh-hold per instance of strategies at the individual level during Phase I

The above table is a handy tool for making choices on which capacity building strategy to deploy to meet the particular requirement of capacity building. An estimated figure is provided for deploying each instance of the strategy and the maximum capacity gain that could be achieved. For a quantitative comparison across strategies, unit cost (per person) per month per strategy has also been computed. The said cost can be used to compare the average cost incurred on a single project resource to build his capacity through each of the mentioned capacity building strategies for a month's duration.

Data presented in the table above shows that according to the cost per instance of a strategy incurred, sending a trainer to live with the team for short duration seem to be the most cost effective strategy to build research capacity in localization teams. However the limited duration and heavy load of the curriculum to be covered in a short while would at best conclude in a capacity gain level of 2, where the

teams through this mode of training would be able to understand the solution for development. However this strategy could generally be deployed over a team that is already going through development of localized software over a fixed time line. Thus teams when faced with a specific problem could raise a concern and an appropriate national training may be provided to them.

However according to the per unit cost calculation of the expenses incurred, the sustained approach to the accomplishment of research goals through development of deliverables stands as the most cost effective solution to train teams and can raise their skill level to maximum of 4, being able to implement the solutions.

Figures indicate that designing a regional training for technical skill enhancement on the other hand is an expensive solution, in terms of cost and the maximum capacity gain gained by the project teams. It is believed that such a mode of training can only raise the capacity of individuals from complete un-awareness of a problem to being able to understand the complexity of the solution to be developed. However it is believed to consume a lot of resources. The positive factor of organizing such training is in terms of networking and building sustainability into the project work which cannot be achieved by any other means. These venues provide an opportunity for the teams to give a face to the names they have been interacting with and would be able to share problems and their intended solutions in a much rapid than through any electronic medium.

Mentor placement program is the second most cost effective solution in terms of the benefit that can be achieved from this strategy. This strategy can promise a maximum capacity gain where the teams would be able to implement solutions and progress towards development of expert commercial solutions. However if this arrangements is cut short on time, then it has an inverse effect on the strategy's effectiveness on average capacity increase.

Mentor placement can however raise the capacity of a certain team only where the training is taking place. However organization of a summer school, though seems to be the most expensive solution in terms of the cost, but can potentially raise the capacity of the group to a level where the trainees are able to progress towards development of expert commercial software. Summer school has an additional advantage of imparting theoretical knowledge that the team's practice during the training and get familiar with the discipline taught.

Based on the experience of the PAN Localization project, numbers have been computed in the table below that provides a starting point for cross- comparison of strategies on salient factors.

Strategy	COST (In 1000 of USD) Per Instance of Strategy	Time Upper Limit	Unit (Person) Cost Per Month For The Strategy	Potential Capacity	Interpretations
Deliverable Development	9.9	36 months	275	4	Able to Implement the Solution
Mentor Placement	7.8	6 months	46	4	Able to Implement the Solution
Regional Training	73	1 week	3685	1	Able to Understand the Problem

Table 3.29 Cost, duration and capacity Gain thresh-hold per instance of strategies at the individual level during Phase II

Based on the experience of phase I, Pan Localization Project organized more mentor placement programs during phase II and it had been successful. All countries included in the program achieved the targeted capacity gain level 4. For instance after mentor placement program Cambodia country team had been successful to submit all required localized software as per the contract during phase. Similarly, after having received training on OCR under mentor placement program, the Nepal and Laos country teams were able to develop and release the relevant OCR software in a timely manner. As far as cost effectiveness is concerned, the success in phase II was even better than in phase I, as in phase II the mentor placement program was found to be the most cost effective one, as in phase I, it stood at second position. The deliverable development which was the most cost effective strategy in phase I, was the second most cost effective strategy in phase II.

3.5.1.4 Support to Present at Workshops and Conferences

As presented above a number of research publications have been produced by the partner countries. This was a testimony of the maturity of their research skills developed during the project. As a strategy for motivating teams to produce more publishable research, RS provided teams with an incentive for covering the conference registration expenses as well as their travel and stay for presenting the research paper at the conference. In Phase I. partner countries participated in 12 conferences and workshops where as in phase II, they participated in 40 conferences and workshops. This strategy significantly motivated teams to produced publishable research.

Strategies Regarding Training to Conduct Close to Practice Research

An outreach component for the project research work was specifically implemented with most of its project partners during the second phase of the project, while the first phase had focused on development of the technology. For this purpose, in the second phase the project also developed partnerships with civil society organizations to specifically focus on dissemination of technology to end users in the partner countries, with explicit funding allocations to support the partnerships. For example, Nepalinux developed by Madan Puraskar Pustakalaya (MPP) was used by E-Network Research and Development (ENRD) to train five rural communities around Nepal, which included farmers, mothers' group, and retired army men. They used the Nepali language applications to communicate with their relatives abroad and to develop online community portals. Similarly, Pakistan component collaborated with District Governments of Sargodha, Chakwal and Attock to deploy localized open source applications in ten rural schools, training more than 200 school students and teachers on information access, communication and content generation.

The partnerships have enabled partners focusing on outreach to appreciate technical challenges and helped the technical partners to appreciate the end user dissemination and adoption challenges. Both lessons significant for planning research they would undertake in the future.

Strategies for the Development of Linkages

As a strategy to develop international collaborations, RS has been organizing regional training, conferences and workshops, in which experts from the region are invited. These have provided opportunities to meet and discuss opportunities for collaboration. As a salient example, project partners

have been interacting with NECTEC, Thailand, which have eventually resulted in formal bi-lateral and multi-lateral partnerships. The project has also worked with researchers from Korea, India, Japan and regional organization like Asian Federation of Natural Language Processing. Such interactions have also resulted in direct partnerships between Microsoft and country partners resulting in the development of LIP for MS Office and Windows in Urdu, Pashto, Bangla, Sinhala, Khmer, and Lao, by the project partner countries.

Strategies to Disseminate Research Work

As a continued effort to enhance country team's capacity to disseminate research work, Regional secretariat has been channeling the regional project funds for research publication and dissemination to country teams for organizing their national events as well. In this regard, the initial Nepalinux and Dzongkha Linux release events were organized through additional funding provided to Nepal country component by diverting centrally available un-spent funds to the country components. Similarly regional marketing budget has been mobilized to bear expenses for organizing country awareness seminars and printing and publication of project promotional materials.

Strategies for the Sustainability of the Research Work

PAN Localization project's regional secretariat had a focused approach on capacity building the teams in ensuring sustainability of the research being conducted. In this context, the project had instituted specific strategies that are detailed below:

Support for Higher Studies

As a strategy for continued and advanced training in local language computing, the project provided completed or partial scholarships for many team members for pursuing higher studies in disciplines related to localization research. Specific researchers were funded in Pakistan, to accomplish their academic research through working on Project. In addition, project facilitated these team members by providing time for studies and examinations and in certain instances by supporting the tuition fee for their degree program. This approach improved the organizational research capacity by having more trained resources in the local language computing domain. While it was also used for retention of researchers, as these team members would remain with the Project until degree completion

Developing Research Centers: to Advance Research Capacity

For sustainability and advancement of localization research, fully equipped research labs were developed, with equipment, specialized resources, including software, books, etc. such that country projects housed within universities and government institutions that would continue the research beyond the project duration. Dedicated research centers were established through the project, for example. In Pakistan, Center for Language Engineering at University of Engineering and Technology was established. This has instigated further localization collaboration and research.

PAN L10n Multilingual Chair in Local Language Computing

To sustain and consolidate the regional momentum of localization research capacity building initiated through the project, a permanent research chair for multilingual computing has been established at the project regional secretariat in Pakistan funded by IDRC, Canada. Establishment of this research chair would provide a sound foundation to sustain, nurture and grow the network of localization researchers, and to provide direct support in language computing community, including researchers and policy makers.

4 ICT Literacy and Local Language Computing

One of the fundamental research objectives of PAN Localization project were to study the impact of localized ICTs in enabling access and use of ICTs for the digitally divided population. Specific research objective included understanding the efficacy of the localized ICTs for enabling online communication as well as access and generation of local language content.

Through PAN Localization project, seven partner countries developed and executed outreach programs to deploy localized ICTs in the grass-root communities across a diverse user groups including students, teachers, farmers, monks, government officials etc. Country partners adopted different models/approaches for disseminating the localized ICTs. For instance, few countries adopted "train the trainers" approach for disseminating the localized ICTs to rural communities. While some focused on providing access to local content to rural communities by using the "mobile lady" model, similar to the mobile tele-centre concept. The countries that deployed the ICTs in rural communities and imparted end-user training were: (i) Bangladesh, (ii) Cambodia, (iii) Nepal, (iv) Pakistan, (v) Bhutan, (vi) Sri Lanka, and (vii) Mongolia. However, for the purpose of evaluation, data was collected from four country partners i.e. Bangladesh, Cambodia, Nepal and Pakistan for determining the impact of localized technology in facilitating access and use of information and communication online.

Surveys were conducted using semi structured questionnaire (attached in Appendix C) to collect data on the project's effectiveness from the end-users participating in the training programs. This survey questionnaire was first developed in English language and later translated into the local language to gather response from the selected training participants. 324 trainees who got training under the PAN localization project participated in that survey, off which, 11 respondents were from Bangladesh, 150 from Cambodia, 102 from Nepal and 61 from Pakistan. These questionnaires were filled by the training participants themselves, about three months after the training program had ended. The data was analysed using SPSS and is presented in the country sections below.

Additional data has also been gathered though secondary analysis of various reports developed by the country partners as well as the regional secretariat. The purpose of using secondary analysis technique was to triangulate the project findings gathered through the survey. These reports were also helpful in providing details about the training methodology followed and reporting country level discussion on the research findings evaluated through various evaluation methods such as survey, pre and post-tests, interview etc. Analysis of OMg frameworks developed by the country partners were also used in identifying the specific boundary partners, associated outcome challenges and progress markers developed for each project. These results can be equally generalized to other developing countries faced with similar conditions, who are striving for expanding ICT literacy in the presence low literacy rates specifically low English language literacy.

This chapter is further organized in two sections. Section 1 discusses country specific model for conducting ICT literacy training and the pre and post assessment regarding the efficacy of the training program in terms of building ICT literacy of the trainees. While Section 2 presents the findings on the language preferences for online information exchange, and impact on, indirectly learning the use of ICTs.

4.1 Country Specific Models for Conducting ICT Literacy Program

Imparting training to indigenous speech communities on the use of localized ICTs was one of the fundamental objectives of the PAN localization project. To meet this objective, each country team followed a unique model or approach to meet this objective. The following section presents these country cases by briefly discussing the models/approach followed and the data on skill enhancement of the end-users to illustrate the efficacy of the training program.

4.1.1 Bangladesh

Through PAN Localization project phase II, D.Net in collaboration with BRAC University developed a rich content database in Bangla language. This effort was part of a larger objective of the organization through which D.Net had been developing Bangla language content for the rural communities following a six phased methodology since 2003. Figure 4.1 below presents the content development methodology followed by D.Net.
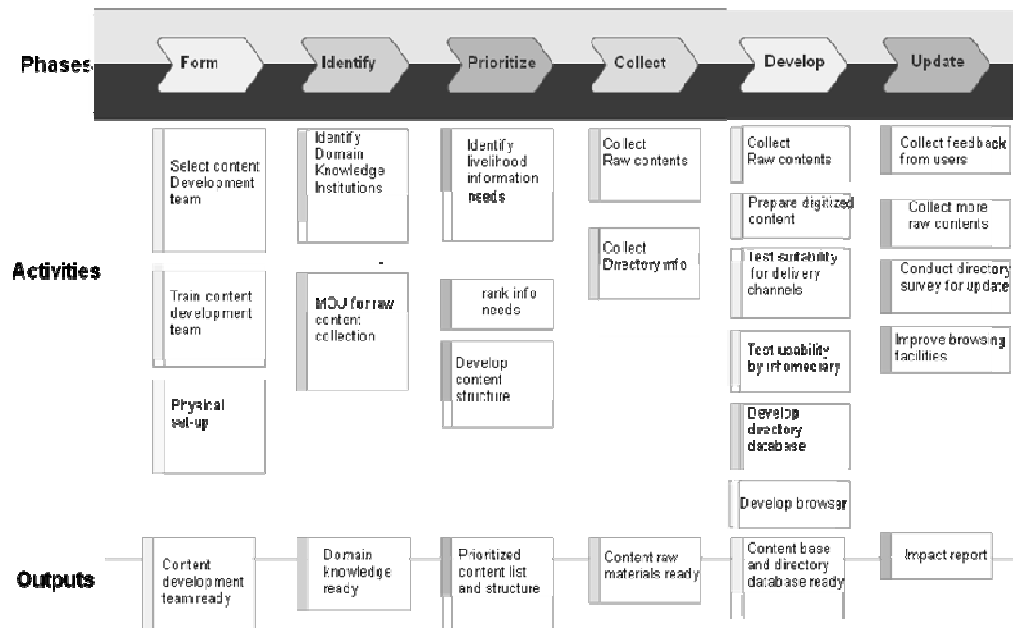
Figure 4.1: Content Development Methodology followed by D.Net

Based on the above approach, in 2003, D.Net conducted a need assessment survey from the rural communities, funded by RIB that identified the need for developing livelihood content for rural community. D.Net then started to develop relevant content based on the identified needs. The content was developed on ten themes of rural livelihood namely, agriculture, education, human rights and legal support, health care, non-farm economic activities, awareness, employment opportunity for rural people,

appropriate technology, disaster management and citizen services. To develop this content, D.Net had been training a group of content developers to compile the relevant livelihood content, by mainly consulting books published by the government and private sector on the given subject. The content being developed was validated by domain expert institutes and expert panels.

Through PAN Localization project phase II (2007-2011), D.Net developed more than 1000 pages on law and human rights content and added this content into Jeon-IKB, the largest Bangla language website on livelihood issues, providing content through both online as well as offline mode (in form of CD-ROM).The content developed during this phase focused on women rights, inheritance issues, land law, family law, local government issues (their structure, services etc.), and domestic and international human right laws. The developed content was made available at the Jeon-IKB subtitled 'Information for Life website, www.jeeon.com.bd.

After the creation of the said content, the next challenge was to enable rural population to use the content effectively by addressing their queries related to rural livelihood through the information. Thus an innovative model for enabling rural access to information was developed by D.NET. D.Net established four *Pallitathya Kendra* (meaning rural information centre in Bangla) (www.pallitathya.org.bd) which is a tele-center established within the rural community to provide rural populations the information they require. Every Pallitathya Kendra has one centre manger, two *informediaries* (one male, one female) and one *mobile lady*

An *informediary* is an information worker who assists the rural community in quickly finding the required information by using ICTs. A *mobile-lady* is like a mobile tele-centre. She moves from door to door with an Eee-PC (Net-book) with internet access and content in the Eee PC, digital camera, and a mobile phone. She may directly search and provide the required content from her Eee PC or respond to the content request through helpline.

Through PAN Localization project Phase II, D.Net conducted training for content developers as well as infomediaries. Content developers were trained to develop locally relevant content while infomediaries were provided training on how to disseminate local language content to rural population. D.Net also conducted refresher trainings to update learning and skill of trainees after six month of trainings. Specific details of the training conducted have been provided in sections below.

4.1.1.1 Training of Content Developers on Bangla Content Development

D.Net conducted training for content developers from 26th - 30th March, 2008. D.Net trained fifteen participants including telecentre practitioners and government officials from different departments such as Dhaka Ahsania Mission, World Net, Dhoritry Samaj Kallyan Shongha, Digonter Dak, Gandhi Ashram Trust, Shariatpur Development Society, Ghashful, Socio Economic Development Association, Participatory Advancement Social Service and Darpan Samaj Unnayan Kendro and Agricultural Information Services, Directorate of Agricultural Marketing on local language content development.

On the first day of the training, content developers focused on report writing, and side by side an overview of the content development in Bangla and existing content development methodologies were

discussed. Training to use content management systems, Wikis and Blogs was also included in this session. Specific training was conducted on the characteristics and layout of Avro, and Joomla. Participants were also trained on writing and posting articles in Joomla based websites. The trainees practiced Joomla by creating its account. All trainees created account in Bangla Wikipedia. They learned how to search topics, edit topic, post articles, insert images, and create external and internal links in Wikipedia. The participants were also trained on creating Blog posts in English and Bangla language.

They were further trained on creating tags to facilitate searching, and raising questions to acquire help. The participants practice its use by creating accounts on "www.somewhereinblog.net" and biggani.com. They also updated their profiles and wrote their views on above-mentioned blogs. On the last day of the training program before certificate distribution ceremony, the participants finalized and shared their strategy plans for content development. D.Net also conducted pre and post evaluation to assess the effectiveness of the training program.

After six months, D.Net conducted refresher training for the content developers to discuss potential solutions for the issues that content developers were facing in uploading content. Nine trainees from different telecentres and government institutes like World Net, Digonter Dak, Gandhi Ashram Trust, Shariatpur Development Society, Participatory Advancement Social Service, Darpan Samaj Unnayan Kendro, Agricultural Information Services and Directorate of Agricultural Marketing attended this training and shared their field experiences.

Figure 4.2: Workshop on Local Language Content Development (25-30 March 2008)

During this session, D.Net revised the curriculum of CMS, Joomla, Wikipedia and Blog. D.Net presented the www.wikimedia.org and desherkhobor.net to raise the interest of the trainees for content development specifically for posting local news at the web. The project coordinator initiated the development of an email group to instantly share their activities in the field.

4.1.1.2 Efficacy of the Training Program

D.Net conducted surveys, before and after the training program, in order to assess the change in knowledge and change in capacity to use the localized software. During the training, a session was

conducted covering the basic concepts of content development including discussions on types of content, the content development life cycle, content development methodologies, content development tools and some principles of effective content development. Later during assessment, the trainees were asked to report their knowledge about various sources of content, different types of content including audio and video content, online and offline content and the steps of content development. The trainees were asked to choose either of the two options; clearly understand the topic, ii) partially understand the topic. The results are shown in the graph below.
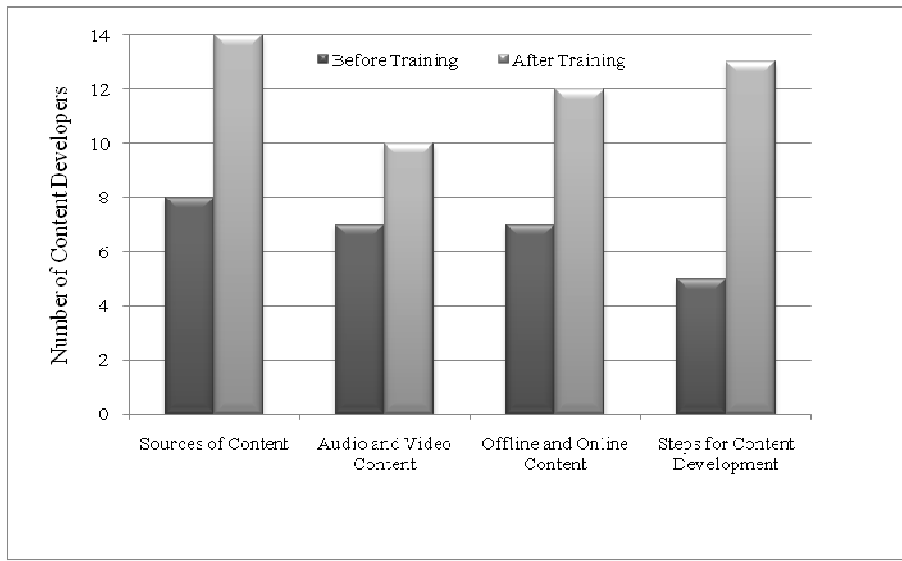
Figure 4.3: Change in Content Developers' Knowledge about Content Development Sources, Types and Steps

When surveyed at the beginning of the training session, out of the fourteen training participants, eight participants knew about various sources of content, seven clearly understood about the differences between various types of content while only five participants were familiar with the steps for content development. However, when surveyed after the training session, all of the fourteen trainees reported that they clearly understood the different sources of acquiring content; eleven reported that they had a clear understanding about the various types of content and thirteen of the trainees reported that they now clearly understand the different steps involved in content development. Thus, the training helped to enhance their knowledge about the content development life cycle and different types of the content.

Another major focus of the training program was to train the participants to use different CMS for content development and search. Three CMS namely Joomla, Wiki and Blogs were introduced in the training sessions. Through pre and post test survey, it was assessed if the participants were knowledgeable about CMS. Further questions were asked to investigate if the sessions had enhanced their prior knowledge by introducing them with new technology. The results of the survey are presented in figure 4 below.
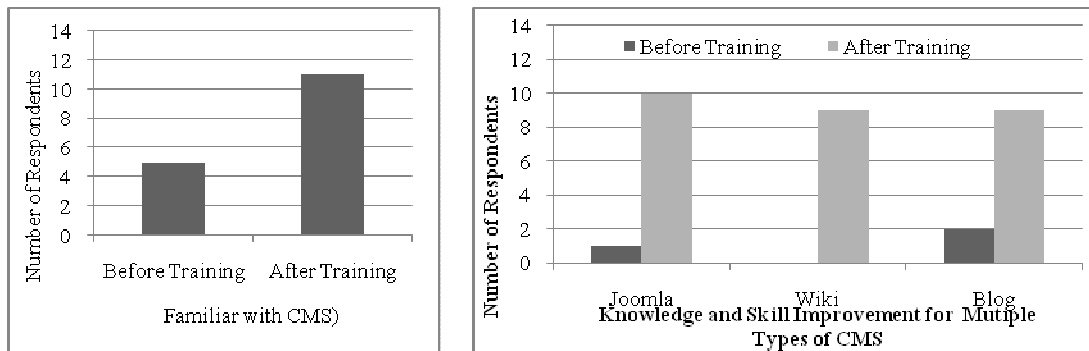
Figure 4.4: Change in Knowledge about Content Management System and Its Types

The findings presented in figure 4.4 above depict that the training helped in improving the knowledge of the participants about various CMS. Results have shown that only 5 participants had some initial understanding of content management systems among whom only one participant was familiar with Joomla and two were familiar with Blog. This is important to mention here that the trainees, who were familiar with CMS, had used English based Joomla and Blog. And, they never used Joomla, wiki and blog in Bangla language. After training, 10 of 14 trainees reported that they were familiar with CMS. During the training, the trainees got hands on experience of using Joomla based website jeeon.com.bd and biggani.com. They also practiced wiki and blog in Bangla. Most of the respondents reported, "*by using Joomla, wiki and blog in Bangla language, content can be developed quickly*" and "*through this we can focus on providing our indigenous knowledge to our future generation*".

4.1.1.3 Training of Infomediaries on Dissemination of Local Language Content

The objective of conducting the ICT literacy training program was also to train the Infomediaries on effective search techniques to be able to find the relevant local language content for the rural population by using computer and mobile phones. This training was organized from 23rd – 31st January 2008. In this training, fourteen participants representing different telecentres/organizations were trained including Rural Knowledge Centre (RKC), Coastal Rural Knowledge Centre (CRKC), Dhaka Ahsania Mission (DAM), Global house of Advanced Technology (GHAT), Rural Technology Centre (RTC) Aramnagar Bazar, E-Hut BRAC Net, WARP, Digonter Dak-D Net.

Topics covered during the training included word processing, spreadsheet, searching the Internet and photo editing. The training also covered tips on effectively searching livelihood content through Jeon-IKB (Off-line and online content database in Bangla) website. A complete session on mobilization and marketing that focused on the importance of enhancing mobilization and communication skills was also conducted. During another session, the trainees were asked to role play as an infomediary and an information seeker from the community. In such a way the trainee acquired practical experience of how infomediaries will be asked to provide the requested information.

Odame (2005) asserts that the existence of ICTs in public spaces does not entail access for all, as on average, women have less income, education, time mobility and face religious or cultural constraints that restrict their access to and use of technology. In this regard, D.Net also included a brief session on

"Gender Awareness" within the content search training of the infomediaries. The purpose of this session was to broaden the vision of infomediaries about ICTs and gender related issues. This session covered concept of gender, construction of gender ideas, the role of gender in development and content, and the gender aspect in the service delivery process. It also covered the barriers that hinder women from receiving and visiting information centers.

A field visit to a nearby *pallitathya kendra* operated by Gandi Ashram Trust in Noakhali District, located in the southern part of Bangladesh was also part of this training. The trainees later visited two villages around the centre and interacted with service recipients of the community. They also observed how the mobile lady worked with the community people and the way she delivered information services and challenges. A few of the villagers also shared their success stories with the trainees. To gauge the effectiveness of the workshop, D.Net also conducted pre and post-evaluation of the workshop.

A refresher training of the Infomediaries was conducted after six months of the initial training. The basic purpose of the refresher training was to discuss and resolve issues faced in operating Pallitathya Kendra. During this training, the infomediaries presented the challenges in using the helpline, offline and online formats of Jeeon-IKB for content dissemination. A special session was also organized during the training on how to operate a telecentre/pallitathya Kendra in which the participants discussed common types of information usually requested by rural community; issues in online and offline content searching; and the most appropriate channels for content delivery.

Figure 4.5: Infomediary Training on Dissemination of Local Language Content

The refresher training also included a session on effective ways of motivating rural community on content access. Three different motivational processes were discussed during the session. They were, organization of seminars, meeting at courtyards and meeting/discussion with *Union Parisad* to attract the rural community members towards telecentres.

4.1.1.4 Efficacy of the Training Program

To measure the efficacy of the training program provided to infomediaries, pre and post surveys were conducted to measure the skill improvement/enhancement of the trainees. The trainees were asked to rate the clarity of their understanding regarding the topics: fundamental concepts related to content and its development, identification of various sources and channels to acquire relevant contents. Based on their self analysis, the following graph presents their responses.

102
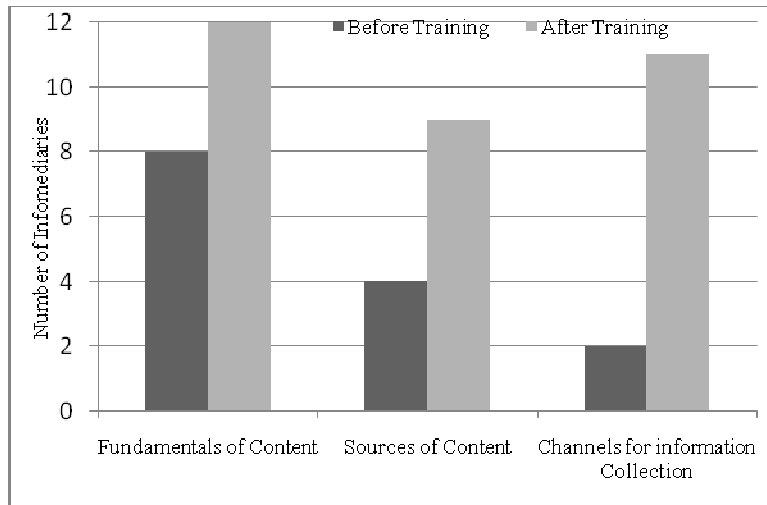
Figure 4.6: Change in Infomediaries' Knowledge about Sources of Content and Channels for Information Collection

A significant change in knowledge about fundamentals of content, sources of content and channels for information collection is observed as shown in figure 4.6 above. The graph shows that eight of the fourteen participating infomediaries were familiar about the fundamentals of content before training while after the training conclusion, all the participants were confident that they were familiar about the fundamental concepts of content and its development. Regarding prior knowledge about sources to search for content only four infomediaries reported they had some understanding in this regard; however after the training, nine infomediaries reported that they now had good understanding of the required sources. Based on the data presented above, it can be concluded that the training significantly helped to raise the learning and knowledge of the training participants regarding content development.

4.1.2   Cambodia

In Cambodia, PAN localization project was carried out by Institute of Technology, MoEYS, and NiDA. The training on Khmer FOSS application was undertaken by NiDA from 2006-08. This initiative was aligned with the national ICT goals and objectives of the Cambodian government. The aim of the "Khmer FOSS Training program" carried out by NiDA was to enhance ICT skills of citizens and to improve their daily living conditions. Through PAN Localization project, this training program was conducted in eight different provinces of the country primarily selected on the basis of availability of electricity. These provinces included: (1) Kampong Speu, (2) Kampong Thom, (3) Kampot, (4) Koh Kong, (5) Preah Vihear, (6) Pursat, (7) Utdor Meanchey, (8) Krong Kep. These training were conducted at Regional Teacher Training Center (RTTC) or a Provincial Teacher Training Center (PTTC), Provincial Vocation Training Center (PTC) or a University that was chosen because of its central location within the province. To conduct the training NiDA was established a fully equipped computer lab at the training center.

Within this training program, NiDA focused on its objective of providing "ICTs for all". Training program adopted the train-the-trainers model. At first stage, NiDA trained its partners. At second stage, NiDA conducted the training and its trained partners assisted the NiDA team. At the third stage, the trained partner conducted the training with the support from NiDA team. In all the trainings NiDA provided training material developed in Khmer language. Participant to computer ratio was ensured to be

103

2:1. Further, NiDA also established a Help Desk to support the trainers and users in their training. Following this training methodology, 1,366 participants were provided hand-on training to use localized applications. The training sessions conducted by NiDA covered the following modules.

Module 1: Linux administration training and certification, Linux administration teachers training including advanced use of user interface

Module 2: Training of trainers who will impart FOSS end-user training

Module 3: Development of training materials in Khmer for Linux Administration and a Linux user interface (KDE)

In Linux administration training, NiDA trained 90 participants from different ministries. The duration of this training was 6 weeks and almost 30 hours per week. For training the trainers to impart FOSS – end user training, NiDA trained 600 participants. The duration for this training was almost 20 hours. The teachers who were familiar with computer and had little expertise before training received training for trainers modules. The teachers who had not used computer before received training on word processing and spreadsheet only.

The FOSS modules included open office (Calc, Impress, Word Processing), use of internet, Mayura, Mekala, Suse, and Unicode. Three provinces, Kampong Speu, Kampong Theu, and Kampot conducted a full suite of training modules. Five provinces, Kokong, Preah Vihear, Pursat, Utdor Meanchey and Krong Kep were given the training only on "Calc".

Figure 4.7: Training In Kampot Province, 3-7 November 2008

The training program included a diverse group of trainees. The training participants included government officials (local official, provincial government official, national officials); the training providers (school teachers, university teachers, vocational training institute, and teacher training center); the school students; the IT industry such as those in the computer shops to take on the role of troubleshooters; the NGOs working on IT applications; and the internet cafe owners. NiDA's objective to train government officials and students was to enable them to use computers for official work and communications. The intention to train training providers was to give them exposure on how to deliver lectures electronically

and generate electronic coursework. The Internet cafe owners were trained to guide cafe visitors and raise awareness in community about FOSS. The purpose to include NGOs in the training program was to help the community to use Community Information Centers and in creating awareness about ICTs. In few provinces, the military and police personnel also participated in FOSS training conducted by NiDA. This is worth mentioning that NiDA assembled this diverse group under one platform and imparted trainings. The figure below presents the training program conducted at Kampot province.

4.1.2.1 Efficacy of Training Program

NiDA used an opinion based survey to see how far was the training program successful in making trainees' able to use ICT is their routine official and personal work. The data was collected from 440 trainees out of 1336 at the end of the training program covering 8 provinces. Among the 440 participants, 68% were government officials, 15% belonged to police sector, 6% of respondents were training providers, 5% were from education sector, 2% were from finance sector and rest belonged to health, interior ministries.

In this regard, when asked to report the extent to which the FOSS modules (open office (Calc, Impress, Word Processing), use of internet, Mayura, Mekala, Suse, and Unicode. 403) had helped them to accomplish their work, 45% of the respondents replied that the Foss module helped them to accomplish their routine tasks and 29% of respondents considered that FOSS applications is "very helpful" in accomplishing their routine tasks. While only 3% of the respondents reported that the FOSS applications helped them a "little" to complete their routine tasks. The data illustrated that training program significantly helped trainees to accomplish their routine tasks as shown in figure 4.8 below.
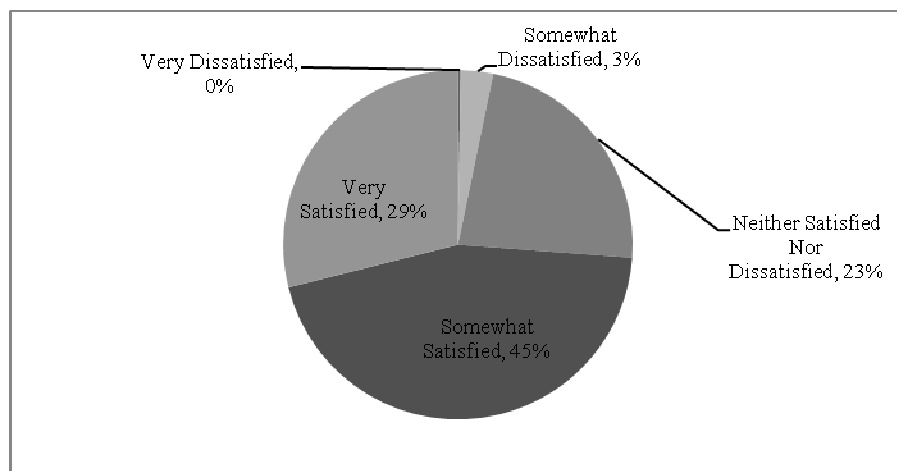
Figure 4.8: FOSS Module Facilitated In Accomplishing Routine Tasks

Similarly, when asked that to what extent the training satisfied their need and expectations, majority of the trainees (211) reported that they were "very satisfied" from the training program and 146 trainees reported that they are "satisfied" from the training program. 58 trainees were neither satisfied nor dissatisfied from the training program. While only 4 trainees reported that their expectations from training were not fulfilled.

According to the team lead of the training program, in order to conduct a FOSS training program successfully, "*Learning objectives and instruction design must match with the curriculum and level of trainee and the trainer. In our target group, the levels of trainees were different. Thus, we adopted a strategy to make the training session more effective in learning aspect. The persons who had little expertise in using computer were allocated the seats with the persons who were not familiar with computer. Therefore, we allowed trainees to take some of the roles of the trainers during training sessions. He or she for example can be made to perform as the class assistant in some of the training sessions. This brought positive outcomes.*" Table 4.1 below presents that the needs and expectations of the trainees were largely fulfilled.

Training Locations (Provinces)		Very Dissatisfied	Somewhat Dissatisfied	Neither Satisfied Nor Dissatisfied	Somewhat Satisfied	Very Satisfied	Total
Sr.	Name						
P 1	Kampong Speu			2	25	15	42
P 2	Kampong Thom			5	21	31	57
P 3	Kampot		2	18	21	18	59
P 4	Koh Kong		1	6	30	8	45
P 5	Preah Vihear		1	7	39	19	66
P 6	Pursat			4	17	22	43
P 7	Utdor Meanchey			3	47	19	69
P 8	Krong Kep			13	11	14	38
Total		0	4	58	211	146	419

Table 4.1: Training Satisfied the Need and Expectation of Trainees

The data illustrated that majority of the trainees were satisfied with the training program. As in Kampong Speu, 25 trainees out 42 reported that their needs and expectations were somewhat satisfied from the training program and in Kampong Thom 31 out of 57 trainees were "very satisfied" from the training program and so forth. However, a significant number of trainees reported that training satisfied their needs and expectation from the training program.

The team lead of the NiDA also shared a strategy which led them to deliver such a successful training, he stated, "*we used and integrated "old wisdom" in the delivery of the training especially for new learners. We created a comfortable environment for new trainers through pep talks. We done this to motivate them and overcome their fear from technology. For example, we said to trainees that tiger and monkey do not talk but rural people always learn, how, because you can communicate. So, why are you scared? We observed a change in confidence among participants after hearing this. When the participants start talking with others, they begin to learn and make the first steps. The motivating statements which we used for the staff and for potential trainers was: (i) don't die with your knowledge; (ii) two heads are better than one; (iii) you have short arm one cannot hold the mountain alone; and (iv) poor can help but lazy cannot help.*"

Based on the findings presented above it can be concluded that the NiDA training program was successful as high percentage of trainees were not only able to use Foss application but also used Foss application in their routine tasks.

4.1.3   Nepal

MPP and ENRD conducted training on localized ICT in two phases. The first phase (April-September 2008) focused on providing training on Nepalinux and other open source software to the end-users. The second phase (October 2008 to March 2009) included advance training of the end-users on creating digital content using localized software.

During the first phase of the training program, training program executed by MPP and ENRD was based on "train the trainers" approach. The training was conducted successfully in three stages. At first stage, telecentre operators from each of five selected locations were trained. Then, each of these five trainers managed LLLT in their respective centers to five other persons who belongs to different groups such as teachers, students, youth etc. Subsequently, these five participants managed the third level training for their same group in each center. MPP and ENRD trained rural community on localized technology in five different locations of the country. The locations included Danda Gaun in Rasuwa district, Jhuwani in Chitwan district, Tolka in Kaski district, Nangi and Shika in Myagdi district. The training program focused on farmers, women, students, youth, and many other groups. Thus, at the completion of the project, 31 end-users's were trained in one center and 155 in all centers. The respective detail about each layer of the training is discussed in the sections below.

The First Layer Training focused on the "TOT", which aimed to develop local trainers in the selected five locations. MPP in collaboration with ENRD trained 5 telecenter operators from 25th November to 4th December, 2007 in Kathmandu. The main objective of the training was to make the participants familiar with localized software including NepaLinux, localized CMS, etc., so that they could easily provide technical support and further training to their community.

Telecenter's Locations	Gender	Farmers	Youth	Women	Students	Teachers	Ex-Army	Hotel Professional	Forest Users	Grand Total
Danda Gaun,	M	1				1				5
	F		1	1	1					
Jhuwani Chitwan	M	1	1		1	1				5
	F			1						
Nangi, Maygdi	M		1			1	1		1	5
	F			1						
Tolka, Kaski	M		1		1	1		1		5
	F			1						
Shikha, Myagdi	M				1	1		1		5
	F		1	1						
Grand Total		2	5	5	4	5	1	2	1	25

Table 4.2: Number of Trainees Participated in Local Level Leader Trainings (LLLT)

The objective of second layer of training was to train local level instructors for further deploying the localized computing technology in rural community. This layer of training was called as LLLT. This training was organized from 22nd January 2008 at four locations while in Tolka Kaski, it started from 28th January 2008. At this stage, every trained teacher/ tele-center operator developed his/her own training outline and course in local language and nominated a group of five participants to whom he would impart

the required training. Each group had representation from target population of women, farmers, students, youth and teachers. Through this training, total 25 participants were trained to use local language software. Among them, 2 trainees belonged to farmers' group, 5 belonged to youth group, 5 belonged to women or mothers' group, 4 were students, 4 were teachers, 1 was an ex-army man, 1 belonged to hotel professionals group and 2 belonged to forest users groups. The complete detail of trainees with respect to gender is given in table below.

At third stage, each LLL trained his/her own community members. This training stage was called End-User level Training or EUL. The objective of training at this level was to provide basic knowledge about the use of different localized open source software and enabling people to use computer in their routine work. This third layer End-User training started from 1st week of May 2008 at all the centers. Again, after receiving the training, each LLL was required to develop his/her own training manual. Also, each local level leader was responsible to impart training to 5 more people belonging to his/her own group. Like farmers conducted the training to 5 more farmers, students conducted the training to 5 more students and so on. However, interestingly almost every LLL not only trained their respected group but also identified and trained new groups as paper making group, Dalit group, and ward committee group. In this training, total 125 (77 female and 48 male) participants were trained on using Nepalinux and other local language software. Among them, 16 trainees belonged to mother/women group, 27 belonged to youth group, 21 trainees belonged to students group, 20 belonged to teachers groups and 21 belonged to farmers group. The complete detail of trainees with respect to gender and group is given in table 4.3 below.

Locations	Gender	Women/ Mother Group	Youth	Students	Teachers	Dalit	Paper Making	Ward Committee	Hotel Professionals	Farmers	Other	Grand Total
Shikha, Myagdi	M			5	2	2						25
	F	5	5		3	3						
Nangi, Maygdi	M				3			2			1	25
	F				2		5	3		5	4	
Tolka, Kaski	M		5	3	1				1	2		20
	F			2	4					2		
Jhuwani Chitwan	M		6	2						6		30
	F	6	6	4								
Danda Gaun, Rasuwa	M		1	1	5							25
	F	5	4	4						5		
Grand Total		16	27	21	20	5	5	5	1	21	5	125

Table 4.3: Number of Trainees Participated in End-User Level (EUL) Trainings

The duration of the phase two of training program was October 2008 to March 2009. During this phase, MPP and ENRD conducted the second training for the participants of the first "TOT" and also organized a workshop to raise awareness about the achievement of PAN localization project.

The Second TOT was conducted by the joint efforts of ENRD and MPP in Nangi Telecenter of Myagdi District from November 9, 2008 to November 18, 2008. The content of this training was more specific than the first TOT. This training was organized for the operators of the five sites who had participated in the first phase training. Nevertheless, it was not only limited to them. The objective of this training was to train participants for developing local language websites by using CMS, blogs, introducing new features of Nepalinux and to teach them about domain registration and Web hosting.

Figure 4.9: Second Training of Trainers (ToT), 9-18 November 2008

Another objective of the second TOT was to involve rural community as well in creating local language content. Thus local community was empowered to create content by the villagers and for the villagers. Following this motivation it was envisioned that students and teachers would be able to put the study materials, farmers would be able to create and accesses agriculture based content and so on. ENRD was also interested to create a local language portal for the local community where they could post their comments about the created content. Thus ENRD and MPP focused on developing blogs for each village. In this context, ENRD and MPP used two approaches for content development. One was top-bottom approach and the other was bottom–up approach. In the first approach, the local language content was produced by ENRD, MPP and other organizations. In the second approach, local content was produced by rural community including teachers, students, villagers and local government. The addresses of the developed websites are: www.shikha.com.np, www.nangi.com.np, www.jhuwani.com.np, www.tolka.com.np, www.dandagaun.com.np. The content posted on these websites includes local language content on educational material, poems, stories, and advertisement of the local products, local news and tourism.

ENRD and MPP also organized a workshop in this phase on 16 March 2009 at Dandagaun telecenter, Rasuwa. The objectives of this workshop were: (i) to discuss about the activities of Dandagaun telecenter in the village; (ii) to highlight the role of PAN Localization training in the development of local knowledge; (iii) to draw attention of the local leaders and policy makers towards the importance of localized ICTs and (iv) to discuss about problems local people are facing while using ICTs. Telecentre operator of Rasuwa, two teachers and two principals of different schools of the Rasuwa village, a representative of Village Development Committee (VDC), and two journalists also participated in this workshop. The trainees who participated in trainings shared their views and experience with other

members of the workshop. They discussed that many villagers were interested to learn localized ICTs. The trainees also expressed their determination to volunteer time for conducting similar training in the village. They also shared some of the limitations and problems that they faced during the trainings. Most of the limitations, that were discussed, were related to the in availability of required infrastructure like computers, etc.

Figure 4.10: Workshop at Dandagaun telecenter, Rasuwa, 16 March 2009

MPP and ENRD trained 155 rural participants including tele-centre operators, students, teachers, women or mothers, farmers, hotel professionals, papers makers etc. To measure the effectiveness of the training program in terms of gaining ICTs skills by using localized ICTs and the enhancement in their ICT skills to train others, ENRD conducted a survey to get the baseline data and then at the end of the training program using "interview" technique to get feedback regarding the benefits of the training program.

4.1.3.1 Effectiveness of the Training Program

A survey was conducted from trainees at the beginning of the training. The purpose of this survey was to determine if the trainees were familiar with the computer prior to the training, and also to gauge the skills they possessed about its use. All participants (155) took part in this survey including 5 trainers, 25 LLL and 125 end-user trainees. As the project was executed in 5 different locations thus each location consisted of 31 participants. The results of this survey are given in table 4.4 below. Table represents the data about the trainees who were familiar with the computer.

The average percentage in table 4.4 below shows that only 38% of trainees were familiar with the computer before training. In Danda Guan, a highest percentage 56% was familiar with computer as compared to other locations. In Shikha village, 43% of trainees were already familiar with computer before training. 40% of the trainees in Tolka, 32% in Nangi and only 20% of the trainees in Jhuwani village were familiar with computer before training.

Name of Locations	% Of Trainees Familiar With Computer Before Training
Shikha, Myagdi	43%
Nangi, Maygdi	32%
Tolka, Kaski	40%

Jhuwani Chitwan	20%
Danda Gaun, Rasuwa	56%
Average Percentage	38% (or is equal to 59 trainees out of 155)

Table 4.4: Percentage of Trainees Familiar with Computer

The familiarity with computer here means that trainees had some exposure about computer however this does not reflect their competency to use the computer. As the Nepal evaluation team reported, "*Among 59 (or 38%) trainees out of 155, only 5 trainees who participated in "TOT" were familiar with the basics of computer and the Internet. The trainees who participated in LLLT had seen and heard a lot about computer but they never got chance to touch or use computer. The trainees who participated in EUL training only heard about the computer and they never got chance to see the computer."*

As, the training program employed "train the trainer's" approach to conduct localized ICT training program for the rural communities, thus, the efficacy of the training program was strongly dependent on the ability of the newly trained participants to successfully train the new trainees and successful development of the curriculum to conduct the training in their own style. As mentioned earlier, the training was successfully conducted in three stages. The following figure 4.11 concisely depicts the efficacy of the training program in the light of above-mentioned explanations.

Stage 3: EUL Training

Stage 2: LLLT

125 trainees who never saw computer before training were using computer successfuly after training

25 Trainees who never used computer before training trained 125 trainees on Localized ICTs succefully after training

Stage 1:

5 Trainees who were only familiar with basics of computer before training trained 25 trainees succesffuly on Localized ICTs

Figure 4.11: Efficacy of Training Program on Localized ICTs

At first stage of the training program or TOT level, 5 participants were trained. Although all the trainees were the basic users of computer, yet they had no knowledge about Open Office Software and NepaLinux. They were trained on Microsoft Windows and Nepalinux operating system installation knowledge (Nepalinux live CD, NepaLinux in PC), Open office and Microsoft office application installation (Open office calc, Open office empress), Open software Base Internet Browser and Messenger, Mozilla Firefox and Gimp Messenger, Troubleshooting and error management (Hardware, Operating System). After training, each trainee successfully performed the role as trainer and trained 25 participants during the second stage of the training namely LLLT. This is worth mentioning that each

participant of TOT developed his/her own curriculum for LLLT. On the whole, almost all trainees of TOT trained the participants of LLLTs on basic computer knowledge specifically knowledge about Open Source Software, Open Office Suite (Calc, Impress), Open software Base Internet Browser and Messenger (Mozilla Firefox and Gimp Messenger etc).

As the Nepal evaluation team reported that, the participants of LLLT had almost no knowledge of computing system. Therefore, this is worth mentioning that after receiving in-depth training from the participants of TOT on OSS in Nepali, they could operate the software system and develop content according to their needs. They were also empowered to provide training within their own professional groups as farmer trained farmers and students trained students. Each participant of LLLT further provided trainings to almost 5 rural community people. Thus, 25 participants of LLLT trained 125 participants during the third stage of training program namely EUL training.

The evaluation team of Nepal reported, "*Before training, all the participants of EUL training (total 125) had no knowledge about ICTs and they didn't have any idea on how to use the computer. They were curious to gain access and use ICT knowledge for daily activities. After training on Nepali Open Source Software, all participants became aware about the computing system and were able to use the computer. After training, they were able to write letters, documents, and send mail. The participants of EUL are now involved in diffusing ICT knowledge and its positive impacts in their local community.*"

Mr. Narayan Prasad Poudel participated in LLLT held in his village. He was born in B.S. 2029 in Lumle Village Development Committee ward no. 9 Landruk of Kaski district. He is a teacher in Landruk Primary School of Landruk which is the same school he had started his education. He passed his S.L.C. exams form Himalaya Secondary School.
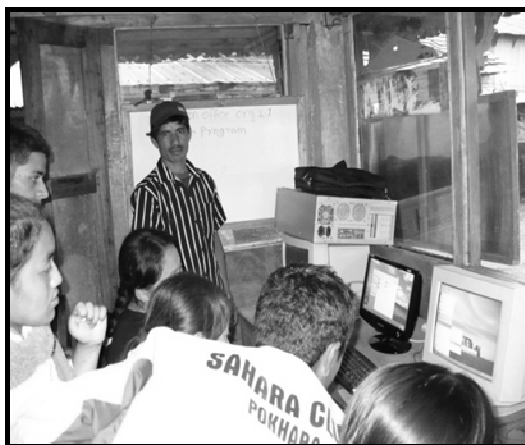
Figure 4.12: Participant of LLLT "Mr. Narayan" Conducted EUL Training

Before the project, he didn't know much about computers. He reported, "*I didn't know how to use computer before PAN Localization training conducted at Tolka telecenter in Nepal*". After completion of LLLT, Narayan said, "*Now I am very much confident about using computers. I can prepare documents; send e-mails, use internet and other things as well.*" Narayan also successfully trained 5 participants of EUL training.

Hence, it can be safely reported that training in Nepal was significantly helpful in enhancing the ICT literacy of the local communities through the use of localized ICTs.

4.1.4 Pakistan

Through PAN localization project Center CRULP, National University of Computer and Emerging Sciences (NUCES) conducted a project named Dareecha to specifically research the following three issues: (i) *effective model for ICT literacy training for rural area school children; (ii) preference for and impact of language on learning the use of ICTs; (iii) differences in ICT adoption across gender for rural area school children (Hussain et al, 2012.).*

The project was conducted in three phases, within a period of three years, each phase lasting for about a year. In the first phase, the ICT literacy training program was designed and all relevant tools and materials were prepared, closely considering the needs of rural area public school children, both boys and girls. In the second phase the training sessions were conducted for participating teachers and students across the participating schools. Detailed project planning was also done in the first phase and careful project monitoring was performed in parallel to the training during the second phase. This planning and data collection was done using the Outcome Mapping framework (Earl et al., 2001) explicitly extended for gender considerations (Shams et al., 2010). In third phase, the project data analysis and evaluation was conducted and the results were synthesized for dissemination to the stakeholders.

During first phase of the project, the key learning areas, curriculum and localized software were designed and developed. The four high level functional areas identified by the project included: (i) General Computer usage, (ii) Information Access, (iii) Communication, (iv) Content Generation. These functional areas were sub-categorized into seven learning areas for ICT literacy. Granular and incremental competency levels were then derived for each learning area. Finally, these competency levels were used to formulate training, practice and assessment materials for the students.

The first functional area "general computer usage" includes basic skills required to use computers for the first time. This functional area had a single learning area: basic computing skills. The second functional area "information access" covered modes of information access available through desktop computers and suitable in a rural school environment for students, and includes a single learning area: web browsing. The third functional area "communication" covers synchronous and asynchronous modes of communication available through computers and useful in a secondary school environment, sub-categorized into two learning areas: emailing and instant messaging. The last and fourth functional area "content generation" contains a variety of learning areas. It includes the following three learning areas: word processing, graphics editing and webpage development.

In order to develop a training program, practice material, and a corresponding ICT literacy evaluation methodology, competency levels were developed for each of the seven key learning areas. Competency levels for each learning area consisted of a set of entry criteria, exit criteria, and a set of levels. The entry criteria defined the pre-requisite knowledge required for a student before embarking on a course based on the levels. The exit criteria defined the level of knowledge a student is expected to have attained after

attending a course designed using the levels. The levels itself broke down the learning area into manageable steps. For example, the competency levels designed for the Basic Computer Skills area consisted of six levels, where the description for the first level is "the student is aware of the concept of a computer and is familiar with the different components of a computer". The description for the last level is "the student is aware of simple problems that may be encountered and is equipped to conduct basic troubleshooting". Each description was then further described in a series of steps. The number of levels developed for each area varied depending on the breadth of the key learning areas.

Software was selected corresponding to each learning area. For the purposes of the program, it was decided that the software should be open source to avoid licensing costs and also because open source software would provide some flexibility in localization options. The following software was selected for each leaning area. For Basic computer skills, Microsoft Windows XP was selected as the operating system. SeaMonkey suite was selected for web browsing (Sea Monkey Navigator), email (SeaMonkey Mail & Newsgroups) and web page development (SeaMonkey Browers). For Word Processing, OpenOffice.org write and for Graphics Editing, OpenOffice.org Draw was selected. All softwares were localized and freely downloadable form the project website.

The training material was also developed in this phase. Training material was developed in Urdu language and customized for students of both genders. Seven training books (one for each learning area) were published by Dareecha team, to be used for the training of both students (girls and boys) and teachers (men and women). The training material was freely distributed to all participating children and teachers through the schools. Though the initial plan was to hand the books over to the students, after discussing with the schools, the books were given to the school library and then checked out to the students for their use. This enabled the schools to get the books back and use them for subsequent batches for training. These books, shown in figure 4.13, are available for free download under the Creative Commons License at from www.cle.org.pk/dareecha.

Figure 4.13: Dareecha Training Books

Assessment material was also developed during this phase in form of ICT literacy tests, which also served as a quantitative evaluation tool to determine technical competency of the training participants. Three types of tests were designed and developed for each learning area as described below:

- Pre-training test: to be taken by students before commencement of the training for a particular learning area. The result of this test was taken as a baseline for the initial competency
- Post-training test (short term): to be taken by students immediately after the end of training on a particular learning area (possibly on the same day). The results from this test were used to measure the extent to which the learning area was learnt by the participants
- Post-training test (long term): to be taken at the beginning of the next training (approximately two months later). The results from this test measured the long term retention of the concepts taught during the training sessions and the effectiveness of the practice sessions coordinated through the teachers at each school during the two months since the training

In the second phase of the project, the training sessions were conducted for teachers and students from 3rd November 2008 to 12 June 2009. The training was conducted in 10 schools including 5 government girls' schools and 5 government boys' schools. These schools were located in the Tehsils of Sarhodha, Bhalwal ,Sahiwal, and Silanwali and Chakwal and Attock. The program planned to train 14 students in each girls and boys schools and two teachers of each school. Thus, it was planned to impart training to 20 teachers (10 male and 10 female) and 140 students (70 male and 70 female). In the schools, where the number of students was almost double, the training sessions were repeated twice daily to accommodate all the students. One of the reasons to accommodate all the students was the overwhelming response of the students specifically girls to learn computer. However, 20 teachers (10 male, 10 female) and 228 students (140 females, 88 male) were trained on localized ICTs.

None of the participating schools had a computer lab. Therefore, computer labs were setup through the Dareecha training program. Lab setup at each of the ten schools included a network of eight refurbished computers (with one setup as a server), a printer, a scanner, speaker's sets and head phones. Special keyboards with Urdu characters printed along the English alphabet were acquired separately and also distributed with the computers, as absence of Urdu letters may have had adverse effects on literacy (as also pointed out by Hall et al. (2009). Internet connectivity was provided through PTCL local loop service. External antennas were installed, as these were remote locations and required the signal to be amplified to enable proper communication. The set up also included furniture for the equipment and to seat twenty people.

The training program was organized and executed in three phases. Each phase was focused on 2-3 learning areas. Phase 1 included two learning areas: Basic Computing and Web Browsing. Training Phase 2 also based on two learning areas: Email and Instant Messaging while training phase three included three learning ares: Word Processing, Graphic Editing and Webpage Development. Each phase started with a teacher training session. Therefore, each teacher worked as a facilitator during students training sessions. The duration of the each training phase was 3 days (8 hours per day) for teachers and 5 days (3 hours per day) for students. The training for teachers was organized at one location while the trainings for students were organized in their respected schools.

Two training teams worked in parallel to conduct the training sessions. The program was implemented in both boys and girls school, and with the social context in mind, a team composed of female members only was formed for the girls' schools, and a team composed of male members only was formed for the boys' schools. The two teams worked together during the teacher training sessions (which included all teachers of the program, both male and female). Each training team had one linguist, who conducted the training, and one technical person, who served as technical support and back-up for conducting training sessions. During this phase of the project, evaluation officer also travelled with training team and collected data qualitative and quantitative data though using multiple data collection techniques mainly survey, interview, observation.

In third phase, the final evaluation data was collected and all monitoring and evaluation based data was analyzed and the results were synthesized for dissemination to the stakeholders. The core approach followed by evaluation team was based on Gendered Outcome Mapping (Shams et al., 2010), also complemented by other methods.

4.1.4.1 Effectiveness of the Training Program

PAN regional evaluation team was interested to know the effectiveness of training program executed in Pakistan in terms of competency to use computer. The regional evaluation team thus conducted a survey with collaboration of Dareecha team. This survey was conducted after few months of the training program. Thus, the results of this survey show the impact of the training program as well. The data was collected from easily accessible trainees from each participated school. In this survey, the trainees were asked to rate their ICT skills on a 5 point scale described as follows:

Skill Level 1 denotes that end users are not familiar with the application; skill Level 2 denotes that end users are aware about application but have never used it; skill Level 3 denotes that end-users are capable to use application but require more training; Skill Level 4 denotes that they are competent to use application with confidence; and skill Level 5 denotes that end-users are expert in using application and even able to train others on it.

The results of this survey are given in figure 4.14 below. The data was collected from 61 (31 male and 30 female) trainees including 46 students and 15 teachers. As the Dareecha team identified four functional areas and within them seven learning areas for conducting training on localized ICTs therefore, the results are also compiled according to their respective skills level in each of the functional and learning areas before and after the training program.
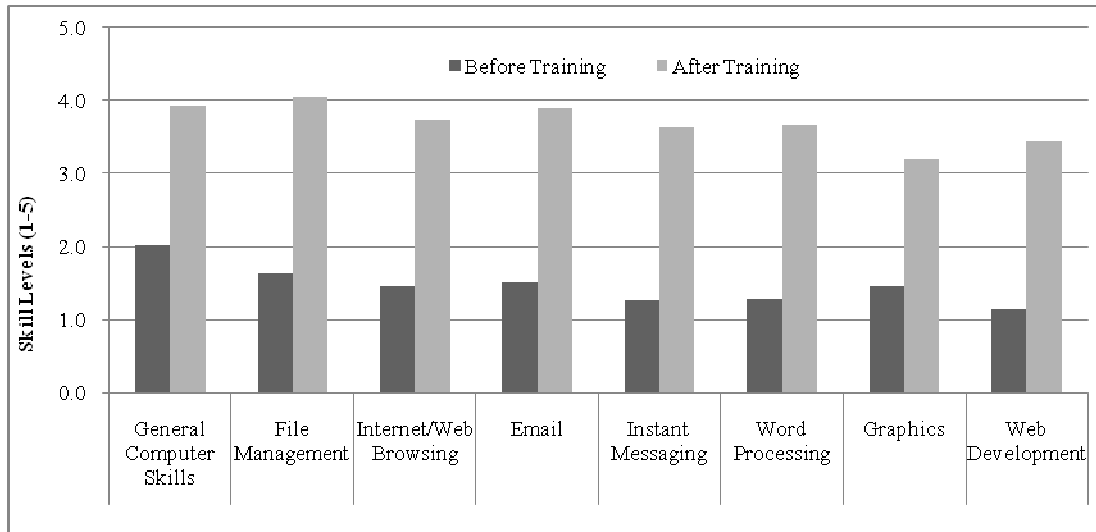
Figure 4.14: Competency in Using Computer (Training Effectiveness)

The figure 4.14 above presents the self assessment of the trainees regarding their competence in each of the functional and learning areas, before and after the training program. The figure presents the "mean" of skill levels. In statistics, "mean" is a measures of central tendency. The graph illustrates that most of the trainee were not familiar with computer before training as the mean skill level for all learning areas except for general computer skill is about 1. The graph depicts that training has significantly helped to raise the capacity of students and teachers to use computer. As most of the trainees reported that after training, they are capable to use software applications but they also need more training. For general computer skill and file management, most of the trainee considered them competent and felt confident in using them.

Similar results were also presented in the country evaluation reports based on the ICT literacy tests taken by the project team to monitor and evaluate the students' competencies over time in various learning areas. The ICT literacy tests were conducted three times during training program: before training (pre), after 5 days training (post-short term) and after almost 2 months of training (post-long term). The following graph shows the results of ICTs literacy test conducted and evaluated by Dareecha evaluation team. Figure 4.15 below presents average score achieved by 139 eighth grade students (70 male and 69 female) in these functional areas.
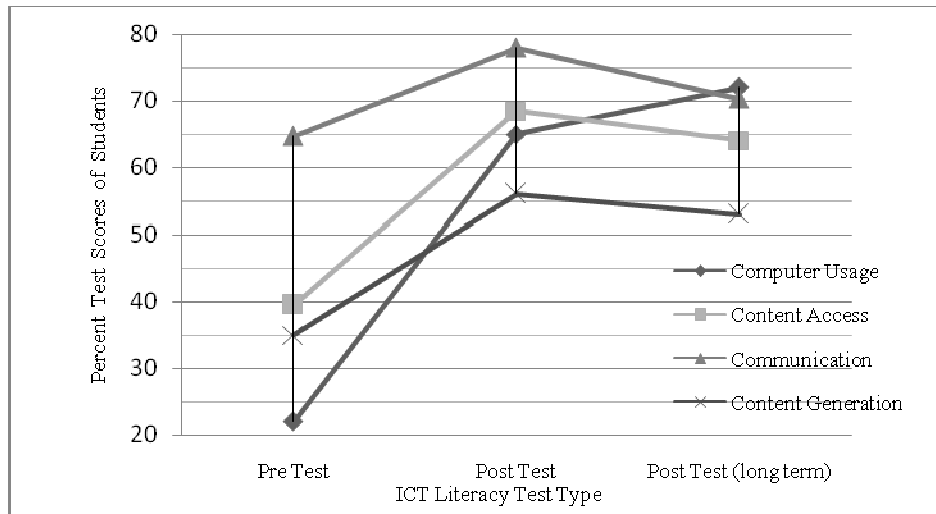
Figure 4.15: Development of ICT Literacy of Students in the Four Functional Areas

The figure above shows that the ICT literacy training helped to raise the capacity of the school students in each of the four learning areas. The figure shows a remarkable change in pre-test scores and post-test scores. This has been observed because the students were not aware of ICT tools beforehand as the above figure 4.15 depicts that after getting 5 days training on respected learning areas, they were not only familiar with the software application but also used them. The most significant increase is observed in computer usage where the results show an improvement from 22% to 72%. Another significant increase is observed in Content Access. The figure illustrates the relatively high scores of pre-training test results for Communication. The country team mentioned a key reason in evaluation report that "*The key reason is that the preceding training on Content Access had already familiarized the students in using web applications, which were also being used for Communication functional area, e.g. web based email and chatting tools, thus raising the baseline.*"

However, the results presented in figure 4.14 and 4.15 are substantiating each other. Both figure shows that training significantly helped to raise the capacity of trainees.

4.2   Impact of Language on Expanding Digital Literacy

Based on the ICT literacy training experiences discussed in the section above, the following section presents specific findings regarding the impact of local language in enhancing ICT literacy of the target communities with a specific focus on role of education and gender upon the findings achieved. In the light of the above, the discussion is specifically focused on highlight the relationship between (i) language and online content access, (ii) language and online content generation, (iii) language and online communication and (iv) language and computer interface. The last section presents the association between education and reading and writing skill of the end-users.

4.2.1   Language and Online Content Access

"The Internet, which has become an important indirect source of information and knowledge and a practical component in education systems, is meant to be for everyone, and then it must speak everyone's language" (ITU, 2010). *To assess the people preferences about which language either English or local language they prefer and usually use for retrieving the online available content on Internet."*

When surveyed, the respondents were asked "In which language do you mostly type search query." Responses to this question were deemed to determine the language in which they preferred to access online content because the search engine displays the searched results showing different web-links in the language in which keyword query is typed. Total 214 trainees responded to this question of which 59 were from Pakistan, 8 from Bangladesh, 81 from Cambodia and 66 from Nepal. The results are given in table below.
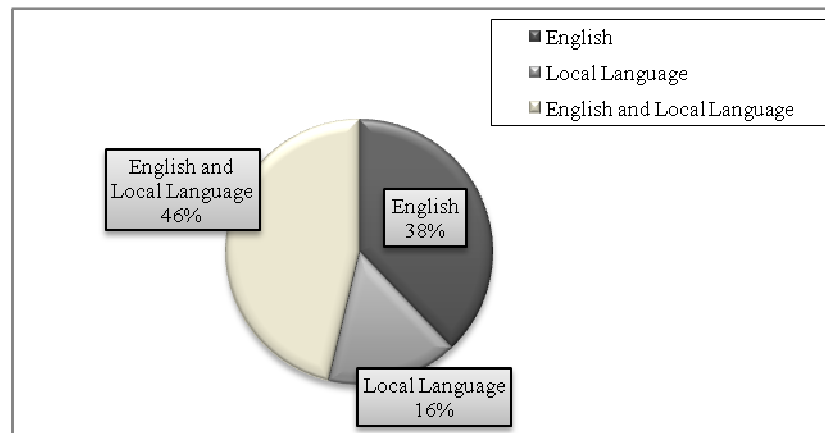
Figure 4.16: Language Preference for Accessing Online Content

The above figure demonstrates that end-users mostly (46%) typed search queries in both "English and Local Language" while 38% of the end-users only typed search queries in "English". The figure also shows that on the whole majority of the end-users typed search queries in English compared to local language.

Similar findings were observed by Pakistan country team while evaluating the student's preferences for accessing online content as reported in the book "Dareecha ICT Training Program for Public Schools in Rural Punjab". Anecdotal evidences unfold one of the reasons of opting for this category of "English and local language". The book affirmed that the current social status of the end-users was the main reason to retrieve the information in both English and local language as "th*e curriculum in participating government schools based on both English and Urdu, that is why most of participated students preferred to access online content in both English and Urdu language"*.

In addition, when inquired from the students, "What sorts of websites have you visited in the last month?", then most of the students 35 (out of 54) reported that they visited "education" related websites. While only 20 of them visited entertainment-based websites, 23 visited news websites, 27 visited religious websites and 34 visited sports websites. On the question which type of information do you

mostly like to search among the categories: (i) work/business, (ii) learning/teaching, (iii) personal needs, (iv) other, majority of the respondents (31 out of 54) reported that they would like to search "learning/teaching" related information. These findings substantiate that one of the reasons of opting "English and local language" category was their current social status.

The above mentioned reason also evokes that education influences people choices to retrieve online information. The table below corroborates that education significantly influence the users' choices to access the online content. Among the 214 respondents who responded to the question above, 6 had primary educational level, 69 had secondary, 56 had higher secondary, 73 had graduation and above, and 8 had other educational levels. The category "other" represents the informal education, technical and vocational education, religious education etc.

Educational Levels	English	Local	English and Local	Total
Primary	0%	1%	2%	3%
Secondary	8%	8%	17%	33%
Higher Secondary	10%	4%	12%	26%
Graduation and Above	18%	2%	14%	34%
Other	1%	1%	1%	4%
Total	38%	16%	46%	100%

Table 4.5: Language and Education Association for Accessing Online Content

The table 4.5 illustrates that the respondents who were on a higher educational continuum preferred to access online content in English while the respondents who were at a lower educational continuum had low preference to retrieve online information solely in English language. As among the respondents who had "graduation and above" educational levels, majority (18% out of 34%) typed search queries in English or in other words preferred to access online content in English language. While the majority of the respondents (17% out of 33%), who belong to secondary educational continuum, prefer to search online information in both "English and local language". The findings also illustrates that as a whole the respondents who had achieved higher educational levels as "graduation and above" usually typed search queries in English.

Through a cross country analysis, it can be found that the majority of respondents who had secondary level of education were from Pakistan, and the reports published by Pakistan country team explain one of the reasons of low preferences to type search queries in English language. The country component team asked the reason for searching the information in Urdu, the students responded that, "*We preferred to search the answer of the query in Urdu language because it is difficult for us to read and understand English language. When we try to search content in English language, we usually do spelling errors.*" The similar response was reported in the Nepal's project progress reports as "*It is easy to search the information in Nepali language as compared to English*".

When the data was analyzed to see any gender differences in accessing the online content, it was noted that male end-users usually typed search queries in English as compared to female respondents. The respondents who answered the question "in which language they mostly typed search queries", among them 150 were male and 64 female. The gender-wise analysis of the same is presented in the figure below.
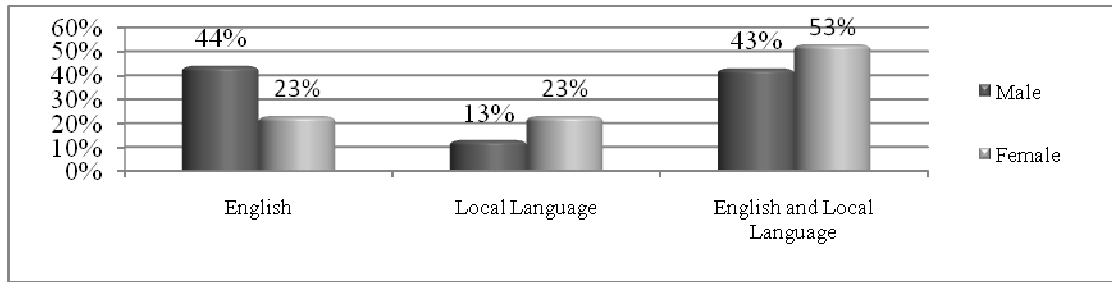
Figure 4.17: Gender Disparity and Language Preference for Accessing Online Content

The figure 4.17 shown above demonstrate that among the male end-users, 44% typed search queries in English, 43% in both "English and Local Language" while only 13% of them typed search queries in "local language" to retrieve online content. Among the female end-users 53% typed search queries in both "English and Local Language" while same percentage equally typed search queries either solely in English or in "local language".

In the light of above presented data about the language preferences for accessing online content, it can be concluded that education and gender play a decisive role in accessing online content as the respondents having higher educational levels highly preferred to access online content in English while the respondents belonging to lower educational continuum have lower preference to search information solely in English language. On the other hand, the majority of the male respondents highly preferred to search information in English while the female respondents have lower preference to search online information in English.

4.2.2 Language and Online Content Generation

One of the objectives of the Pan Localization project was to develop content in local languages using both top-down and bottom-up approaches. By using top-down approaches, the content was developed by the country teams themselves. For developing using bottom-up approach, the country partners provided trainings to end-users on developing online content specifically on how to develop websites and blogs. Among the countries who participated in this evaluation study, two followed bottom-up approach as Pakistan trained the participated end-users to develop websites and Bangladesh provided training to participated end-users to develop websites as well as blogs.

When surveyed that "have you developed any websites", if so, "in which language you have developed websites". The end-users were asked to choose among the categories: (i) local language, (ii) English and (iii) local language and English. There were total 36 end-users who responded to this question, among them 33 were from Pakistan and only 3 were from Bangladesh. The results of the survey are given in figure below.
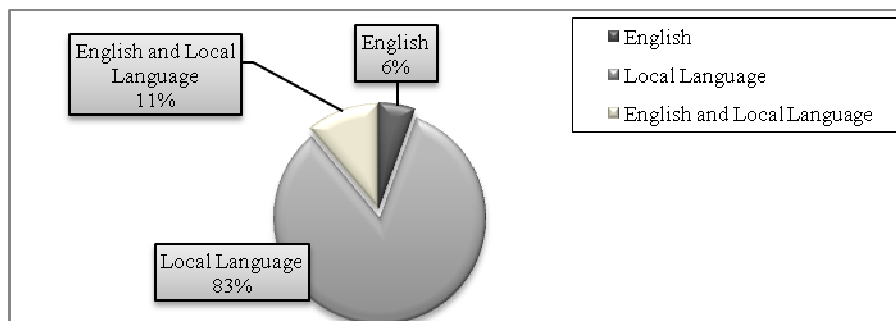
Figure 4.18: Language and Website Development

As shown in the figure 4.18 above that majority of the respondents (83%) developed their websites in "local languages and only 6% of the respondents developed websites in English language. The figures also demonstrates that only 11% of end-users or respondents developed websites in both English and local languages. Similar findings were observed by the Pakistan component who organized a website development competition. The end-users were requested to develop website in any language. Pakistan component received 56 websites of which only one website was in English and 55 websites were developed in Urdu language (local language). When, the evaluation team of Pakistan component inquired the reasons of developing website in local language, majority of the students remarked "*It is convenient for us to produce content in Urdu language because we easily understand it and do fewer mistakes as compared to English language.*"

Gender	Educational Levels	English	Local Language	English and Local Language	Total
Male	Primary	0%	0%	0%	0%
	Secondary	0%	47%	13%	60%
	Higher Secondary	0%	0%	0%	0%
	Graduation and	7%	13%	0%	20%
	Other	0%	20%	0%	20%
	Total	7%	80%	13%	100%
Female	Primary	0%	0%	0%	0%
	Secondary	0%	71%	5%	76%
	Higher Secondary	0%	0%	0%	0%
	Graduation and	5%	5%	5%	14%
	Other	0%	10%	0%	10%
	Total	5%	86%	10%	100%
Average % of Totals		6%	83%	11%	100%

Table 4.6: Language and Education Association for Developing Websites

The above mentioned remarks indicate that the competency to write in a particular language is an essential factor to develop online content. Competency to write a particular language improves or enhanced through either formal or informal education. The data presented in table above shows the association among gender, education and language preference for developing websites.

The table 4.6 presented above illustrates that among the respondents who belongs to secondary level of education did not develop a single website in English language. There were total 68% of the respondents

belongs to secondary level of education of which 60% were male and 76% were female respondents. All the respondents who developed websites in English language had "graduation and above" education. Thus, the findings reveal that education significantly influences the people preferences about using a particular language for website development.

Regarding the question "In which language you have developed your blog", 7 end-users responded to this question. All end-users were from Bangladesh and belong to "graduation and above" educational continuum. The results are given in figure below.
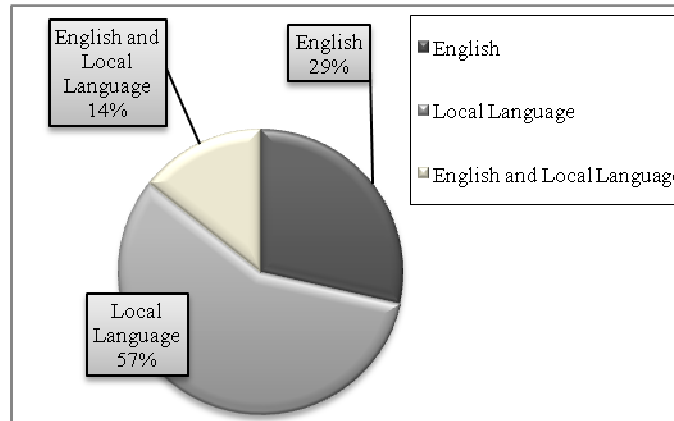
Figure 4.19: Language Preference for Blogging

The figure 4.19 presented above shows that majority of the respondents develop blog in local language which indicates that even it is convenient for the highly qualified person to create content in local language. This is noteworthy that 5 respondents (out of 7) created their own blogs after the training and all these blogs were in Bangla language as: i) http://prothom-aloblog.com/, ii) http://www.somewhereinblog.net/blog/tanzin2008/28783455, iii) http://www.prothomalo.com/ and iv) somewhereinblog.net/blog/kibriyarafiblog. Among them most of the respondents reported that "*they sent new material on a post at least once in a month*".

4.2.3   Language and Online Communication

Internet is one of the cheapest and fast ways to connect with the people and online communication is increasing day by day. For the purpose of this study, online communication was divided into three types: (i) receiving emails, (ii) sending emails and (iii) online instant messaging (chatting). The results are given one by one in sections below.

4.2.3.1 Language Preference for Receiving Emails

The respondents were asked "In which language do you prefer to receive emails?" They were asked to choose between the categories: (i) English and (ii) Other (please specify). Total 201 end-users responded to this question of which 10 were from Bangladesh, 80 from Cambodia, 53 from Nepal and 58 were from Pakistan. The results of the survey are given in figure below.
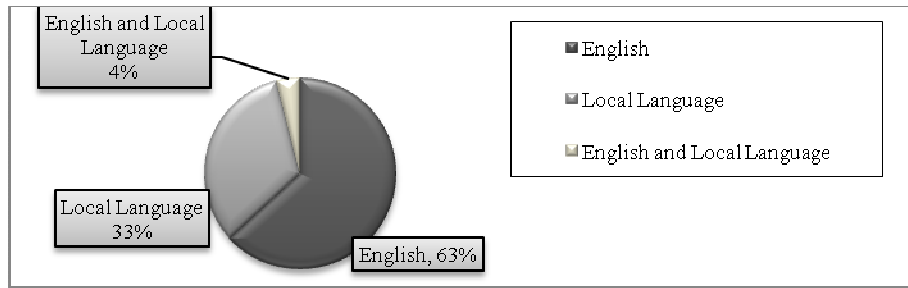
Figure 4.20: Language Preference for Receiving Emails

Figure 4.20 presented above demonstrate that significant percentage (63%) of the respondents preferred to receive emails solely in English language. Yet, also a good percentage of the respondents (33%) preferred to receive emails in "local language". The results shows that the respondents were interested to receive and read emails in solely either in English or local language as only 4% of the respondents opted the category for receiving emails in both "English and local languages".

To measure the educational influence on selecting a particular language for receiving emails, the data is analyzed and presented in table below. Among 201 respondents, 4 had primary, 64 had secondary, 46 had higher secondary, 78 had graduation and above, and 7 respondents had others educational background. 2 respondents did not mention their educational levels.

Educational Levels	English	Local Language	English and Local	Total
Primary	1%	1%	0%	2%
Secondary	10%	23%	0%	32%
Higher Secondary	19%	4%	1%	23%
Graduation and Above	32%	4%	4%	39%
Other	2%	2%	0%	4%
Total	63%	33%	4%	100%

Table 4.7: Language and Education Association for Receiving Emails

The data presented in above table 4.7 shows the extent to which education influences the preferences of people about receiving emails in a particular language. The table illustrates that among the respondents who had "graduation and above" education majority of the them (32% out of 39%) preferred to receive emails in English while among the respondents who had "secondary" level education highly preferred (23% out of 32%) to receive emails in "local language".

To measure the extent to which being male or female makes a difference in language preference for receiving emails; the data has been analyzed and presented in figure below. Among the 201 respondents, 143 were male and 58 were female respondents.
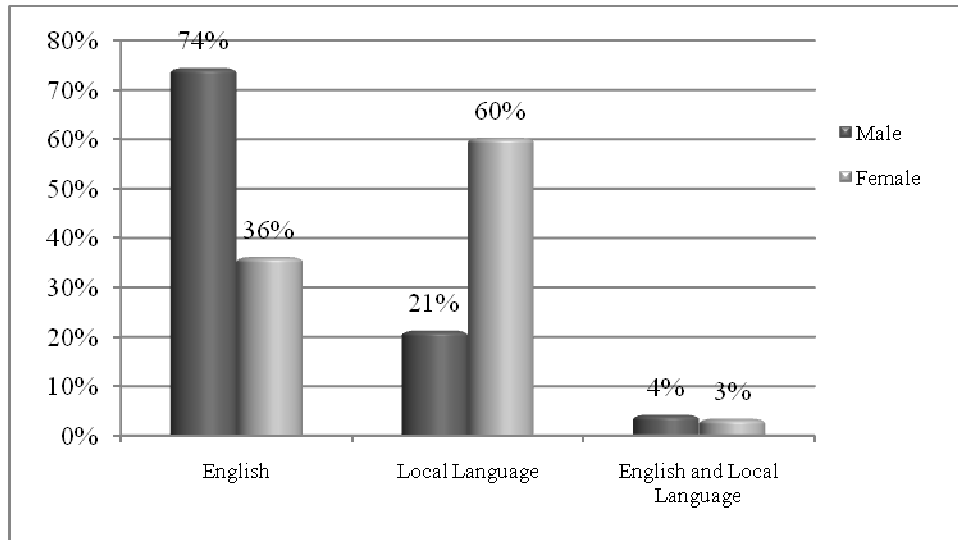
Figure 4.21: Gender-Disparity in Language Preference for Receiving Emails

The above-presented figure 4.21 presents noteworthy gender-differences about the language preferences for receiving emails. As the figure shows that the majority of the male respondents (74%) highly preferred to receive emails in English, while a significant percentage (60%) of the female respondents preferred to receive emails in "local language".

4.2.3.2 Language Preference for Writing Email

The respondents were further inquired that "In which language do you mostly write your emails?", and asked to choose among the categories: (i) English, (ii) Local language, (iii), English and local language and (iv) Local language (Roman script). There were total 175 end-users who responded to this question. Among them 9 were from Bangladesh, 53 from Cambodia, 53 from Nepal and 60 from Pakistan. The results of the survey are presented in pie-chart below.
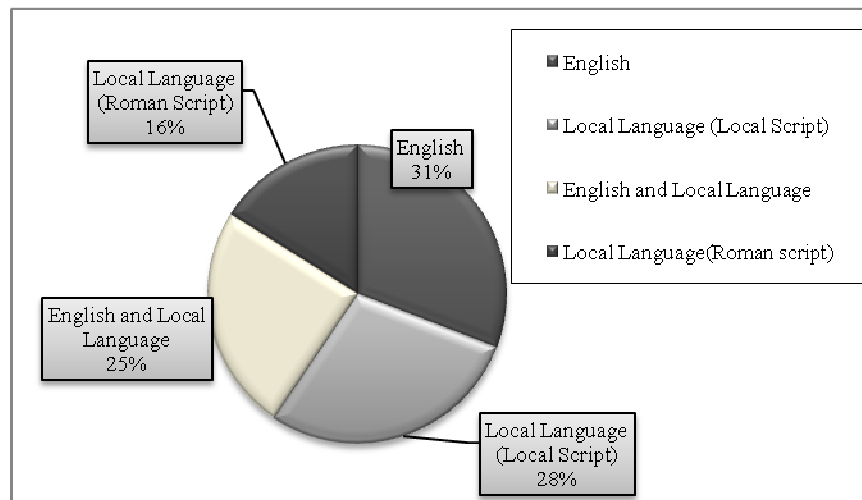
Figure 4.22: Language Preference for Writing Emails

The figure 4.22 presented above shows that majority of the respondents (31%) wrote and sent emails in English language. While, 28% of the respondents preferred to send emails in local language and alike percentage (25%) wrote and sent emails in both "English and local language". This is noteworthy that 16% of the respondents sent emails in "local language (Roman Script).

The above presenting diverse findings shows that people demand linguistic diversity for online communication and this demand is also influenced by various factors. One of the influencing factors is education. To assess the relationship between the education and language preferences, the data is being presenting in table below. There were total 175 respondents of which 2 did not mention their educational levels. Among the rest of the respondents, 4 had primary, 66 had secondary, 38 had higher secondary, 58 had graduation and above, and 7 respondents had other education.

Educational Levels	English	Local Language (Local Script)	English and Local Language	Local Language (Roman Script)	Total
Primary	0%	0%	0%	2%	2%
Secondary	3%	22%	5%	9%	38%
Higher Secondary	10%	2%	6%	4%	22%
Graduation and	16%	3%	14%	1%	34%
Other	2%	2%	1%	0%	4%
Total	31%	28%	25%	16%	100%

Table 4.8: Language and Education Association for Writing Emails

The data presented in above table 4.8 shows the association between the education and language preferences for sending emails. The table shows that the respondents who had higher educational levels significantly preferred (16% out of 31%) to write and send emails in English language while the respondents who had less education highly preferred (22% out of 28%) to send email in local languages. The table also shows that the respondents belong to secondary level of education also prefer to write text of the email in "Roman script".

Another variable "gender" which also influences the language preferences for writing emails is observed as presented in the following figure. Among the end-users who responded to the question about the language preference for writing emails, 117 were male and 58 were female.
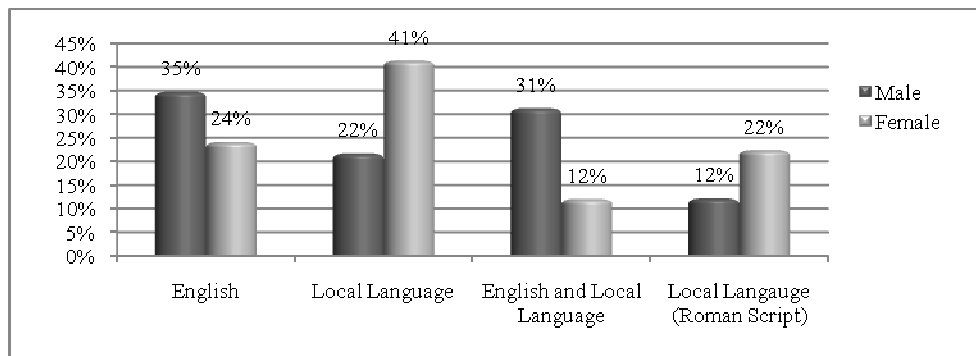
Figure 4.23: Gender Disparity in Language Preference for Writing Emails

Figure 4.23 shown above demonstrate notable gender-difference in the language preferences in writing and sending emails. The figure shows that significant percentage of the women respondents (41%) wrote and sent emails in local language while majority of the men respondents (35%) sent emails in English language. This is noteworthy that female respondents are more interested to write in Roman script compared to male respondents.

4.2.3.3 Language Preference for Online Instant Messaging

To gauge the end-users language preferences for online communication through instant messaging, the respondents were asked to mention "In which language are you more comfortable to chat?" The respondents were requested to choose among the categories: (i) English, (ii) Local language, (iii), English and local language and (iv) Local language (Roman script). The results of the survey from 88 respondents, of which 7 were form Bangladesh, 20 were from Cambodia and 61 were from the Pakistan, are given in figure below.
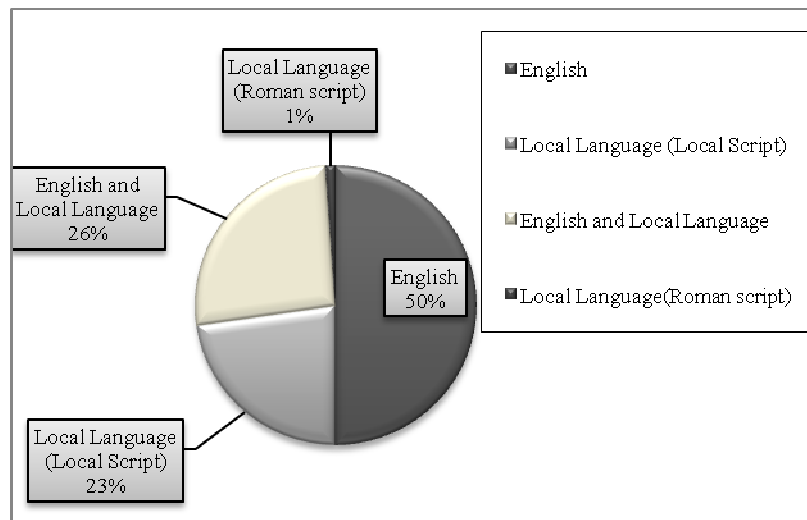
Figure 4.24: Language Preference for Online Instant Messaging

The data presented in the figure 4.24 shown above present that English is frequently used for online chatting by the majority of the respondents compared to local language. As the figure shows that 50% of the respondents chat in English while only 23% of the respondents preferred to chat in local languages. It is widely recognized that people are generally more comfortable in using local language in "Roman script" compared to English or local language. But the data presented in above pie-chart shows that only 1% of the respondents used local language in Roman script for online instant messaging. One of the reasons of this low preference to write the message in Roman script is may be the provision of the software or keyboard to write in local language as most of the respondents were not familiar that the computer can be in local language.

Anecdotal evidence substantiate that the provision of the software or keyboard in local language has shifted the end-users preferences about writing local language from Roman script to local script. As, in Nepal component, the majority of the trainees were those who were familiar with computer before training and used to communicate online in Roman script. It is also observed that few of them were quite

127

happy after learning to type Nepali in local script. As Hem Purja from Nangi village said *"The most important thing that I like about the telecenter in our village is that I can chat freely with my friends and relatives in Nepali. In the past, we used to chat by writing Nepali in English Roman letters but now it can be done in Nepali itself and Nepali typing is also can be learnt in about half an hour."*

Similar results were also observed in Pakistan. During an evaluation study, the project team logged all the written communication of the project staff with the trainees by noting the language in which the communication was done. Total 1467 emails and 363 chats were logged by Dareecha team from Nov. 2008 till Dec. 2009. The results clearly showed the preference of Urdu for communication purposes as 89% of the total online communication was in Urdu, 9% in English and 3% in other local languages (Punjabi, Saraiki etc).

It is also observed that education is an intervening variable that influences the language preference for online instant messaging. To illustrate this relationship, the data is cross-tabulated and presented in table below.

Educational Levels	English	Local Language (Local Script)	English and Local Language	Local Language (Roman Script)	Total
Primary	0%	0%	0%	0%	0%
Secondary	31%	11%	8%	1%	52%
Higher	3%	1%	2%	0%	7%
Graduation and	13%	6%	15%	0%	33%
Other	2%	5%	1%	0%	8%
Total	49%	23%	26%	1%	100%

Table 4.9: Language and Education Association for Online Chatting

The table 4.9 shows the relationship between education and the language preferences for online instant messaging. The table shows interesting results as the respondents who were on lower education levels highly preferred to chat in English compared to the respondent on higher educational level as "graduation and above". These results are giving the dissimilar dimension compared to other online communication findings as presented in section table 7 and 8 above. However, one similar finding can also be observed that among the respondents who usually chat in local language, majority of them (11% out of 23%) belong to lower educational levels as "secondary".

Another intervening variable "gender" was observed during the data analysis and to gauge the relationship between the gender and language preference for online instant messaging or chatting, the data is being presenting in figure below. There were total 88 respondents who reported their language preferences for online instant messaging, among them 54 were male and 34 were female.
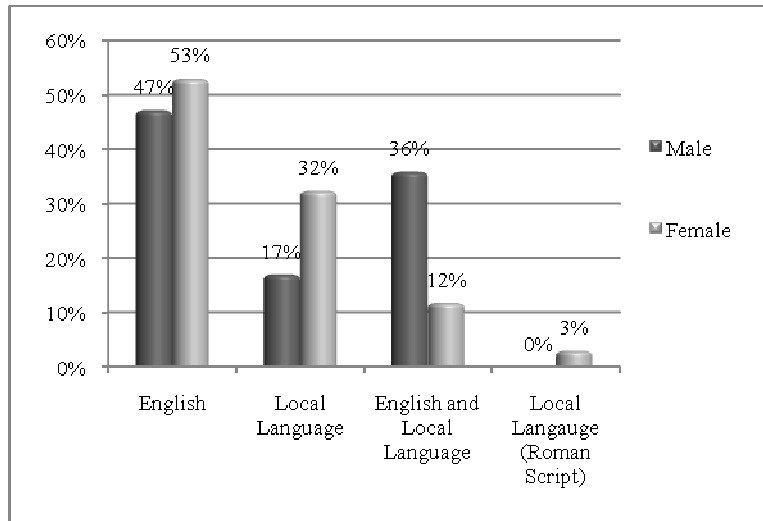
Figure 4.25: Gender Dimension in Language Preference for Online Instant Messaging

The above figure 4.25 illustrates that the majority of the female respondents (53%) frequently used English language for online chatting as well as majority of the female respondents (32%) also preferred to chat solely in local language compared to male respondents. While, a significant percentage of the male respondents (47%) preferred to chat in English language as well as the majority of the male respondents opted "English and local language" for chatting. Thus, the results suggest that on the whole English is significantly preferred by the male respondents compared to female respondents.

4.2.4   Language Preference for Computer Interface

The respondents were asked "In which language do you prefer to use the computer? (i) English interface, (ii) local language interface, (iii) computer in both "English or local language" interfaces? There were total 296 end-users who responded to this question of which 9 were from Bangladesh, 129 from Cambodia, 100 from Nepal and 58 from Pakistan. The findings of the survey are presented in the figure below.
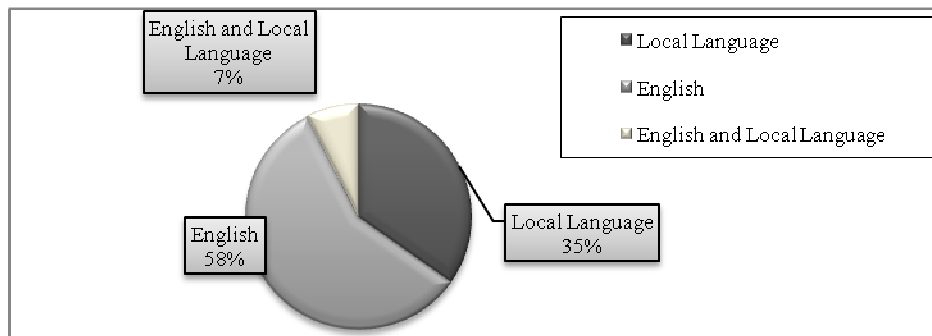
Figure 4.26: Language Preference for Computer Interface

The data presented in the figure 4.26 above demonstrates that majority of the respondents (58%) preferred to use computer with English interface but also a significant percentage (35%) preferred to use computer with local language interface. While only 7% of the respondents preferred to use computer in both

129

"English and local language" interfaces. Thus, the findings suggest that the interface of the computer should be available in both English and local language so the people can use it according to their desire or preferences.

In addition to the findings presented above the data also present the relationship between the education and the language preference for software. This is shown in the table below.

Educational Levels	English	Local Language (Local Script)	English and Local	Total
Primary	0%	2%	1%	3%
Secondary	11%	19%	2%	32%
Higher Secondary	17%	6%	2%	26%
Graduation and Above	29%	6%	2%	37%
Other	1%	2%	0%	3%
Total	49%	35%	7%	100%

Table 4.10: Language and Education Association for Computer Interface

The above-presented table 4.10 shows that among the respondents who had education "graduation and above" majority of them (29% out of 37%) preferred to use computer with English interfaces. While among the respondents who had "secondary" education significantly preferred (19% out of 32%) to use computer in local language interface. Thus, the findings suggests that the education significantly influence people's choices specifically about using the computer with English or local language interface.

Gender analyses of the same data-set reveal that, gender is another significant variable observed in selecting a particular language for computer interface. The results of this analysis are presented in figure below.
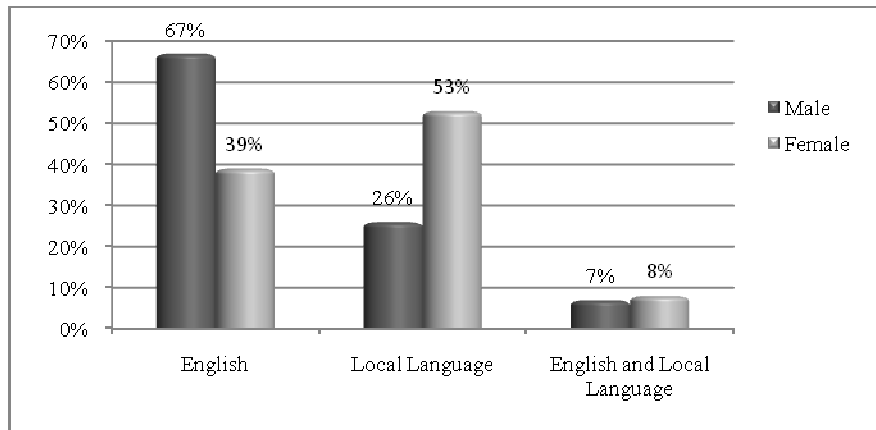
Figure 4.27: Language Preference for Computer Interface

The figure 4.27 presented above demonstrates noteworthy gender-differences in the preference to use computer in a particular language. The figure shows that majority of the male respondents (67%) highly preferred to use computer with English interface while significant percentage of the female respondents (53%) highly preferred to use computer with local language. This difference suggests that it is easy or convenient for female end-users to use computer with local language interface compared to English interface.

130

4.2.5   Reading and Writing Skill in English and Local Language

Online information exchange or learning ICTs requires the competency to read and write English which is lingua franca of ICTs. Though it is widely recognized that people are more competent in mother tongue as compared to any foreign language, yet a questionnaire was still conducted to get some quantitative measure of this difference, specifically to relate competency disparity with educational levels and also to relate it with the results of language preferences for online communication, content accessing and generation. The respondents were asked to rate their skill or competency levels for reading and writing in English and mother/local language. The respondents were asked to rate their skill level among (i) excellent, (ii) very good, (iii) good, (iv) fair and (v) poor. There were total 263 end-users who responded to this question. Among them, 11 respondents were from Bangladesh, 150 from Cambodia and 102 from Nepal. The results of the survey are given in figure below.
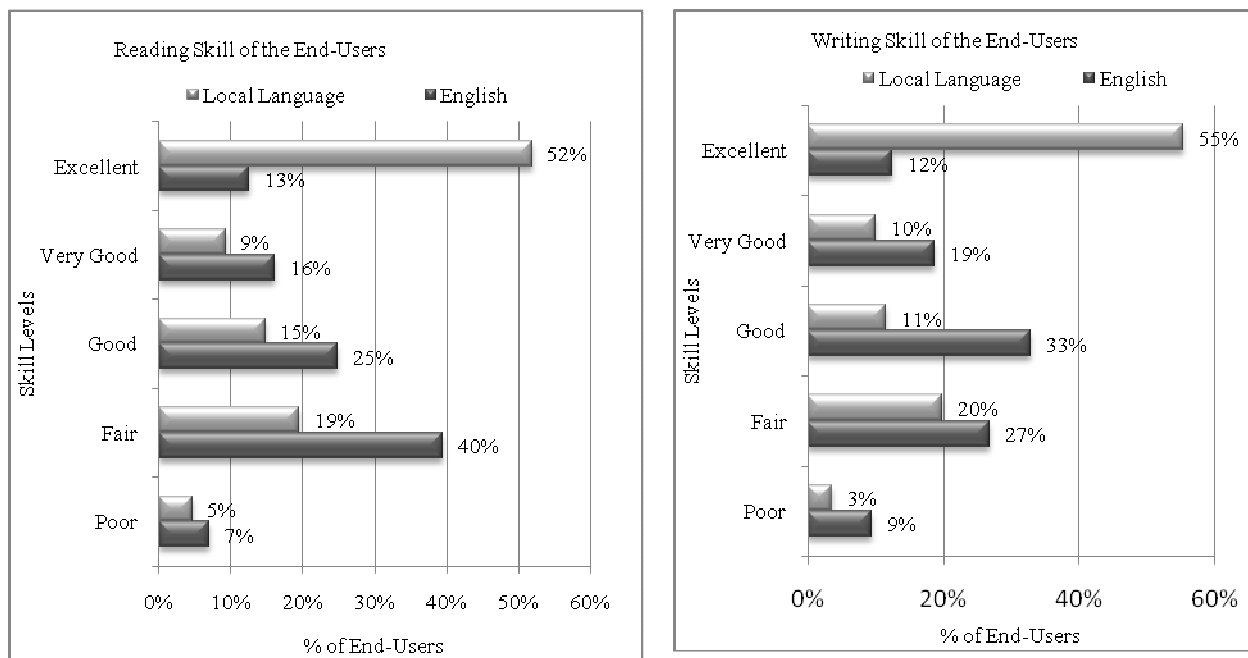
Figure 4.28: Reading and Writing Skill of the Respondents

The figure 4.28 shown in above based on two clustered bar-charts, the bar-chart presented at right show the results of the writing skill of the end-users while the bar-charts at left demonstrates the reading skill of the end-users. The figure illustrates that significant percentage of the respondents (52%) were "excellent" in reading as well as alike percentage (55%) were excellent in writing local language. While, on the other hand a significant percentage considered them as fair in reading (40%) and writing (27%) English language. This is also noteworthy that on the whole only 8% of the respondents were rated them as poor in reading and writing "local language" while double percentage (16%) considered them as poor in reading and writing "English".

The data is cross-tabulated to assess the extent to which the reading and writing skill of the respondents influenced by educational levels. The results are given in table 4.11 below.

Educational Level	Excellent		Very Good		Good		Fair		Poor		Total
	Eng.	Local Lang.	Eng.	Local Lang.	Eng.	Local Lang.	Eng.	Local Lang.	Eng.	Local Lang.	
Primary	0%	2%	0%	0%	0%	1%	0%	1%	1%	0%	6%
Secondary	0%	1%	0%	0%	2%	1%	5%	6%	1%	1%	17%
Higher Secondary	0%	6%	1%	1%	3%	1%	6%	3%	1%	1%	22%
Graduation and	5%	14%	7%	7%	10%	4%	5%	2%	1%	0%	54%
Total	5%	23%	9%	9%	14%	7%	17%	11%	4%	2%	100%

Table 4.11: Education and Readign and Writing Skill of the End-Users

Table above shows the relationship between educational levels and self-assessed competency levels in reading and writing English and local language. The findings depicts that majority of the respondents (23%) consider them excellent in local language as compared to English (6%) language. In other words, the majority of the Bangladeshi, Cambodian and Nepali respondents' consider them excellent in Bangla, Khmer and Nepali respectively as compared to English. The data illustrates that level of education affects the competency in reading and writing English language. As majority of university respondents consider them "very good" in English while the majority of higher secondary and secondary respondents consider them "fair" in English. This is also noteworthy that the respondents belong to secondary level of education consider them "poor" in both reading and writing English while zero percent of them consider poor in local language. The table 4.11 also demonstrates that within the competency levels of English, the highly preferred competency level for English is "fair".

5 Impact of Local Language Computing Research on Policy

PAN Localization project has been a pioneering initiative to conduct local language computing research in south Asia. One of the objectives of the project has been to present recommendations for advancing local language computing policy in the project countries. In this regard, various activities were carried out by the participating countries for initiating debate to develop effective policy for local language technology, training and content within the respective countries.

A summative evaluation study based on the framework for discerning policy influences for strategic evaluation of IDRC supported research projects (Lindquist, 2003) was conducted to appraise specific activities conducted by the country teams to influence country policies, in order to learn key contribution of PANL10n project in developing and influencing local language computing policies.

The sub-evaluation questions developed were:

1. To assess the project's contribution (both direct or indirect) in affecting public policies, laws, regulations, programs and government structures
2. To assess the project's contribution in the development of the network of technologists, social scientist and policy makers for collaborative work
3. To assess how far has the project been able to build the capacity of policy makers on IPR (intellectual property rights) issues

5.1 Evaluation Framework for Discerning Policy Influence of PAN Localization Project

In order to understand the contribution of project teams in developing and influencing local language computing policy in their countries, framework for discerning policy influence for strategic evaluation of IDRC supported research projects (Lindquist, 2001) has been adopted. This framework provides three dimensions for reviewing the policy actions undertaken on a specific issue. These three dimensions are:

1. *Expanding policy capacities*; defined as capacity building and enhancing the knowledge of concerned actors,
2. *Broadening policy horizons;* defined as creating the opportunities to bridge the gap between researchers and policy makers, and
3. *Affecting policy regimes;* defined as an explicit contribution that affects the policy environment.

These dimensions are further explained in the section below.

5.1.1.1 Expanding Policy Capacities

It is found necessary to inform policy makers about latest findings of the research so that they can effectively utilize the information in making decisions (Lindquist, 2001). Because of this gap, the policy makers remain aloof of the research developments and researchers are unable to communicate their

findings in a timely manner in order to influence the policy at the right time. Another important dimension in this context is the *development of talent within the organization for conducting such research and analysis* (ibid). Such capacity building initiatives can help in better communication for policy influence. Thus in case of the PAN Localization project, an investigation was needed to see the efforts of country project team towards influencing the policy makers by improving their knowledge, and informing the policy decisions.

5.1.1.2 Broadening Policy Horizons

Networking provides opportunity to learn and share ideas with others and helps in building consensus on issues related to policy. In addition, such type of networking allows researchers in introducing new concepts and points for stimulation of public debate (Lindquist, 2001). It can be done at national, regional or international level to trigger required change. Through PAN Localization project specific focus was given on networking for dissemination of the research on local language computing.

5.1.1.3 Affecting Policy Regimes

According to Lindquist (2001), "*Research can sometimes influence public policy in a direct way*". Research findings can be used in the development of laws, regulations, programs, or structures. However, such process is rare and normally circuitous, and it happens in fewer cases where change can be attributed, visibly and directly due to the inspiration of research alone. The development of any program and policy or the modification of existing policy is the prime objective that any research project aims to achieve through their findings. Thus country specific policy initiatives of the project teams have been studied to assess if any of the project findings have directly impacted the local language computing policy initiated further research in the knowledge area or has initiated a change in the policy development process.

Based on the policy dimension discussed above, following criteria was developed by adapting Lindquist (2001) approach for the assessment of PAN Localization in influencing the policy environment.

Expanding Policy Capacities	Improving the knowledge or data of certain actors
	Enhancing organizational capacity to conduct policy relevant research
Broadening Policy	Providing policy makers with opportunities for networking and learning
Affecting Policy Regimes	Initiating local language computing projects nationally or internationally
	Affecting implicit contributions in policy decisions/ decision making process
	Undergoing changes in policy development process (inclusion or exclusion of
	Translation of project outputs into policy decisions/ laws/ regulation directly

Table 5.1: Three Dimensions of Policy Influence Lindquist (2001)

In order to assess the contribution of the partner countries in influencing country policies data was collected from the project leaders through a structured questionnaire (attached in appendix D). In addition, policy relevant information from the interim final reports of PAN Localization developed by the regional secretariat and annual/research reports from partner institutions were also analyzed. Moreover, interview and informal discussion were also conducted with representative of each CPI to gather information. However for the purpose of evaluation in the subject report, data has been collected from

seven country teams that were DIT, Bhutan, MoEYS, Cambodia, NAST, Laos, InfoCon Co. Ltd, Mongolia, MPP, Nepal, UCSC, Sri Lanka and CRULP, Pakistan.

This chapter is further organized in two sections. The first section presents the country case studies by defining the political context of the policy environment of the country. It further highlights the key policy making organizations and relevant provisions in the policy documents facilitating incorporation of local language computing clauses. Following the description of the environmental factors, each case study then presents the project achievements in influencing policy as per the indicators defined in table 5.1 above. The second section summarizes the discussion and concludes by synthesizing the overall contribution of the project in influencing policy environment at the regional level.

5.2 Country Findings

5.2.1 Bhutan

5.2.1.1 Overview of the Policy Environment

Bhutan 2020 (Planning Commission, 1999) and Bhutan ICT Policy and Strategy (MoIC, 2004) are key statements defining the political context for the project. The Government's long-term vision and objectives have been set out in its Bhutan 2020 – A Vision for Peace, Prosperity and Happiness, published by the Planning Commission in 1999. The development of national language – Dzongkha has been identified as a priority in this document. It states that: "*Dzongkha has been a particularly powerful force for unifying the Kingdom, establishing a lingua franca among diverse ethnic groups. It is our national language and we must seek to ensure that the position it occupies is further reinforced*" (Planning Commission, 1999). According to Bhutan 2020, the Dzongkha Development Commission should actively promote the use of Dzongkha and ensure that it is responsive to the future as well as the past.

BIPS launched in July 2004, serves as the main policy statement on the development of ICTs. The overall policy objectives of BIPS initiatives are:

1. Use ICTs for good governance
2. Create a Bhutanese info-culture
3. Create a 'high-tech habitat

Under each objective 'Content and Applications' strategy has been identified with the following targets:

1. Establishing the framework for e-business, using ICT to preserve Bhutan's cultural heritage
2. Enhancing the quality and accessibility of health and education
3. Broadening national media and web presence
4. Supporting good governance

BIPS lists under the 'content and applications' initiatives the development of 'a digital archive of significant Bhutanese religious texts and cultural contents in sound and picture format'. Department of

Information Technology (www.dit.gov.bt) DIT started National Digital Library Project in 1997 to provide texts, audio, videos, and images on Bhutanese culture and development related issues.

5.2.1.2 Impact of PAN Localization Project's Research on Local Policy

In Bhutan, PAN Localization project was conducted by DIT, MoIC. DIT is the leading department working under for the development and coordination of ICT-related activities in the country.

In the context of expanding policy capacities specifically by improving the knowledge of concerned policy makers in Bhutan on local language computing issues, the project team held regular meetings with relevant policy making organization to present their research. Six monthly meeting on the project's progress were held at DIT, where other officials from DIT and Dzongkha Localization Project team reviewed the progress achieved by PAN Localization project and pointed out the need for future research. Specifically, in one of these meetings the need to launch version 2 of Dzongkha Linux distribution was identified as a national priority and decisions were reached to promote it for public use. The other factor specifying the need to expand policy capacity is the organization's capacity to deliver policy relevant recommendations. When questioned regarding the organizational capacity of DIT to conduct policy relevant research, the project leaders rated it as high (4) on a scale of 1 (very low)- 5 (very high) points.

Efforts on broadening the policy horizon, has been evaluated in terms of the opportunities provided to policy makers for networking and learning about the project's research. Thus at the completion of project milestones, DIT conducted formal project's output release ceremonies inviting the policy makers, media personnel and the general public. During these ceremonies while the project research were presented, the software developed would also be freely disseminated, e.g. specifically during the release ceremony of Dzongkha Linux 3.0, over 3000 copies of Dzongkha Linux 3.0 CDs and keyboards with Dzongkha Layout have been freely distributed. Government officials have also been regularly invited in the trainings, seminars and other activities conducted by DIT.

To evaluate the impact of research work in affecting policy regimes, initiations of new projects regarding localization research both nationally and internationally have been analyzed. In this regard, a significant contribution of the project has been the development of a MoU between DIT and NECTEC. This MoU specifically focuses on the promotion of R&D in the area of ICTs, specifically in language computing and included a plan to strengthen open source natural language processing, image and speech processing technology in Bhutan. Another significant contribution of the project has been in the acceptance of the research work for the development of local computing standards thus directly impacting/creating a policy. This is because DIT is the lead department working under MoIC for the development and coordination of all ICT-related activities in the country, thus the work done under the project directly contributed towards local language computing policy in the country. Research work accomplished by the project team on terminology translation for Dzongkha Linux with the support of Dzongkha Development Commission has been standardized. DIT also collaborated with cc TLD that operated .bt domain to create the test bed for deploying Dzongkha IDNs.

5.2.2   Cambodia

5.2.2.1 Overview of the Policy Environment

In 2003, recommendations for a national ICT policy were developed within a study of the ICT sector performed by NiDA with the assistance of UNDP-APDIP (Unger & Robinson, 2007). The scope of this study included policy, infrastructure and access, content and applications development, human capacity development, and ICT enterprise development. Since then, it has been under review by the Council of Ministers for several years. The following excerpt from the draft of national ICT policy (Draft http://www.unapcict.org/ecohub/resources/cambodias-national-ict-policy-draft) highlight the key provisions provided in the draft that focus on localization or local language computing. They are as follows:

Policy on Human Capacity Development

3.4 *"Promote access to ICTs nationwide, by adopting a two pronged approach, firstly, it will promote and support general access methods like the Community Information Centers (CIC). Secondly, it will standardize the Khmer script for widespread access to ICT across Cambodia. The Government will mandate NiDA to produce and encourage the production of computer applications in Khmer language, in order to facilitate adoption of ICT in education and in all sectors of society."*

Policy on Standards, Software and Content

4.2 "*Develop and adopt a standard Unicode-based Khmer characters system to be adopted in all ICT applications and systems. Other aspects of this important issue such as keyboard, lexicon, sorting, etc. will also be standardized and codified. The Government will take international ownership of all aspects of Khmer (in Unicode, etc) such that this can be freely used and not be liable for payments for its use in any form."*

4.3 *"Use local language in ICT applications for the rapid development of ICT and at the same time continues encouraging the learning of the English language and its application in ICT."*

4.4 *"Actively promote the development of relevant Cambodian content in all media, including all computer-based systems and networks. It will take the lead in creating all Government websites in a bilingual (Khmer/English) format for the rapid assimilation and use in all sectors from the government to citizen."*

4.14 *"Establish ICT procurement policies aimed at assuring that applications purchased or developed by the government are secured, are localized to the Khmer language whenever possible, and are the best available value, and can be managed by local company. FOSS should be considered as an option."*

In 2005, the MOEYS released their "Policy and Strategies on ICT in Education in Cambodia (http://planipolis.iiep.unesco.org/upload/Cambodia/Cambodia_Policies_strategies_ICT.pdf)".   The following extracts from Policy and strategies on ICT in the education document are the main provisions in it highlighting the government's focus to promote and use local language.

3.2.3 Development of Local Contents, and Use of Open Source Software

b) *"Promote the development of content in Khmer for both formal schools and distance education and equivalency programmes by providing an annual award for the best teaching and learning resources submitted to the National Clearing House."*

c) *"Promote the development of the Operating System in the Khmer language, which will greatly facilitate all literate people in Cambodia to use ICT for communication and learning."*

3.2.5 Knowledge Management System

a) *"Set up an online National Clearing House for teaching and learning resources in Khmer and foreign languages."*

3.2.7 ICT Use in Universities

"Since Cambodia is in serious shortage of well-trained lecturers and professors, it is essential that the universities must use its scarce human resources to their full capacities through ICT as well as connecting learners to virtual learning resources worldwide. The Ministry will promote the use of ICT in teaching and learning process, research and administration by creating a cyber campus consortium and linking this to other virtual universities in other countries. It will also promote the digitizing of the Khmer language books and translating of foreign core books into Khmer."

5.2.2.2 Impact of PAN Localization Project's Research on Local Policy

In the context of expanding policy capacities specifically by improving the knowledge of concerned policy makers in Cambodia on local language computing issues, the project team held regular meetings with relevant policy making organization to present their research. A notable example in this context is the organization of a training session for the Ministry of Interior (MOI) and National Assembly (NA) on utilization of Khmer Unicode and applications package developed through the project's research. This effort made significant development in bridging the gap between policy makers and researchers. The other factor specifying the need to expand policy capacity is the organization's capacity to deliver policy relevant recommendations. When questioned regarding the organizational capacity of MoEYS to conduct policy relevant research, the project leaders rated it as Satisfactory (3) on a scale of 1 (very low)- 5 (very high) points.

Efforts on broadening the policy horizon, has been evaluated in terms of the opportunities provided to policy makers for networking and learning about the project's research. In this regard, Cambodia country component conducted several seminars to introduce its research work to the policy makers. These seminars were attended by a large number of participants from academia and public sectors. CDs containing the research outputs, (Khmer Smart Typing, Encoding Conversion Utilities, Collation and Sorting Utilities, Word Wrapping Utilities, Spell Checker Utilities) have also been freely distributed at these seminars. The project team has also been discussing the possibilities of using the localized application developed through the PAN Localization project in other similar research projects. As a result, Khmer applications were utilized to teach students and the communities at 20 hubs in iREACH project

being conducted by IDRC. Consequently in 2010, through the training sessions conducted by NiDA across many provinces of Cambodia, IT representatives from 24 provinces and cities decided to use Khmer based applications in all districts of the country. Further, as a result of this advocacy, NGO sector has also started to use localized applications in their routine tasks.

To evaluate the impact of research work in affecting policy regimes, the most significant contribution of the project would be the translation of project's outputs into policy decisions/ laws/ or regulation. Thus a significant contribution of the project has been in the adoption of the localized applications developed through the PAN Localization having in Cambodian government. The National Election Committee (NEC) of Cambodia has been using the sorting application developed through the project to sort over six million registration names in Communal Election 2007 and National Election 2008. The Ministry of Finance has been using these localized applications since 2006. National Institute of Language also tested, approved and has been using Khmer based application since its start. Extraordinary Chambers in the Courts of Cambodia (ECCC) used localized applications and accepted Kep font size 12 as a regulation for all official communication.

Another significant contribution of the project is in the initiation of new projects regarding localization research both nationally and internationally. In this regard, a significant contribution of the project has been the development of IDRC's flagship project on the Informatics for Rural Empowerment and Community Health (i-REACH) project. This project has been approved for implementation by the public sector organizations after a significant effort undertaken through PAN Localization project's awareness raising and capacity building efforts. This projects aims to pilot test by offering ICT services to local communities in two areas, Kep and Kamchai Mear. The ICT services provided include access to the Internet and training to use computers within the rural communities. The project participants include government officials, students and monks.

5.2.3   Laos

5.2.3.1 Overview of the Policy Environment

National   Policy   on   Information   and   Communications   Technology http://www.unapcict.org/ecohub/resources/lao-pdr-ict-policy) serves as the main statement on the promotion of ICTs and local language computing to determine the political context for the project. Within its nine priority areas, this national policy identifies "Standardization and Localization" as one of the key focus areas. The Standardization and Localization section of the policy focuses on concrete strategies for promoting localization. The following excerpt from the policy document highlight the key provisions provided for promotion of localization in this document.

Standardization and Localization

"*The adoption of new technology necessarily requires the adoption of national and international standards to ensure network interoperability within and without the country; the preservation of language and culture; and increased productivity through rapid adaptation.*

(i) The GOL (Government of Laos) shall promote software, hardware, and protocol standards, including telecommunications equipment and services, to ensure interoperability and harmonization with international, regional, and sub-regional standards.

(ii) To promote digital interchange in the Lao language, the GOL shall adopt the Unicode standard for the Lao script.

(iii) The GOL shall establish a network, including national and international experts, academia, government, and the private sector to advise on all issues relating to the localization of ICTs – this include Open Source and proprietary software.

(iv) The GOL shall establish a set of standards for data and information storage, exchange, and access for all government bodies and agencies, in line with international data standards. The GOL shall adopt Open Standards for all government data/information systems over proprietary standards, wherever possible, to minimize technological lock-in (http://seacoop.eu/files/2010/03/ICT_policies-programmes_priorities_SEA.pdf).

5.2.3.2 Impact of PAN Localization Project's Research on Local Policy

In the context of expanding policy capacities specifically by improving the knowledge of concerned policy makers in Laos on local language computing issues, the project team regularly interacted and conducted meetings with the policy makers to present their work and get feedback. A specific meeting in this regard was held on January 10, 2008, which was attended by the president of NAST in addition to other policy makers and IT researchers. During this session, the project's research was presented and received a lot of appreciation from the President.

The other factor specifying the need to expand policy capacity is the organization's capacity to deliver policy relevant recommendations. When questioned regarding the organizational capacity of NAST to conduct policy relevant research, the project leaders rated it as Satisfactory (4) on a scale of 1 (very limited)- 5 (very high) points.

Efforts on broadening the policy horizon, has been evaluated in terms of the opportunities provided to policy makers for networking and learning about the project's research. In this regard, the project team presented the project's research and development at the National ICT fair which was inaugurated by the Prime Minister of Laos and had a significant participation of public and private sector. In addition, a technical workshop on the local language computing research conducted by the project team was conducted for 80 government officials, students and journalists, where CD containing fonts, keyboard utilities and open office applications were distributed. The regional conference on "Localized ICT Development & Dissemination across Asia" organized at Vientiane; Laos from 11-16 January, 2009 also provided a good opportunity for Lao team to create awareness among the policy makers about the work being done under the project. The country team also developed its local language website to highlight the project's research (http://www.laol10n.info.la/).

To evaluate the impact of research work in affecting policy regimes, the research work conducted through the project has been directly impacting the local policy. In this regard, Lao Unicode set developed through the project has been adopted by the government of Laos as the Lao National character set.

To evaluate the impact of research work in affecting policy regimes, initiations of new projects regarding localization research both nationally and internationally have also been analyzed. In this regard, a significant contribution of the project has been the collaboration between Microsoft and NAST for the localization of Windows Vista and Office 2007 in Lao. Through this initiative, the government developed a vision to deploy Microsoft solutions across all government departments.

5.2.4   Mongolia

5.2.4.1 Overview of the Policy Environment

In Mongolia, key institutions and organizations contributing to development of local language computing policy are ICTPA, CRC and NITP.

ICT Vision 2010 (http://www.infocon.mn/english/reference/ict_projects/a3.html.) aptly lays out the political context for conducting local language computing research and development in Mongolia. In 2000, Mongolia adopted the ICT Vision 2010 as a blueprint for ICT development in the country. ICT Vision 2010 has three major components: Government-Legislation framework, Business-Economy framework and People-Society framework.

Within the Government-Legislation framework, one of the strategic objectives is creation of policy and regulatory regime. This strategic objective emphasizes the implementation of the following activity related to localization, i.e. *"Ensure the state use of software application developed in Mongolian language."*

Within the Business-Economy framework, one of the strategic objectives is the expansion of ICT business and making it highly efficient economic sector. This strategic objective emphasizes the implementation of the following activity related to localization, i.e. *"Develop market for software applications in Mongolian language."*

5.2.4.2 Impact of PAN Localization Project's Research on Local Policy

In Mongolia, PAN Localization project was executed by Infocon Co. Ltd. in collaboration with teams at NUM and MUST. Infocon Co.Ltd is a consulting company, offering information and communication technology services to government and non-government organizations.

In the context of expanding policy capacities specifically by improving the knowledge of concerned policy makers in Mongolia on local language computing issues, the project team held regular meetings with with ICTPA, Communications Regulatory Committee, Datacom Co., Ltd, Magicnet Co.Ltd and National Information Technology Park to present the project's research. These meetings deliberated on open source software localization, standardization of terminology, IDNs and other related issues. The other factor specifying the need to expand policy capacity is the organization's capacity to deliver policy

relevant recommendations. When questioned regarding the organizational capacity of DIT to conduct policy relevant research, the project leaders rated it as high (4) on a scale of 1 (very low)- 5 (very high) points.

Efforts on broadening the policy horizon, has been evaluated in terms of the opportunities provided to policy makers for networking and learning about the project's research. InfoCon and National University of Mongolia have been disseminating project research and information through their websites as well. InfoCon website at www.infocon.mn, and NUM through the website of the Center for Research in Language Processing developed through PAN Localization project at http://crlp.num.edu.mn.

To evaluate the impact of research work in affecting policy regimes, significant contribution of the project's regular interaction with the key policy makers has resulted in the standardization of the IT terminology translation by ICTPA committee. Moreover, discussion with Datacom (.MN registrar) on localized IDN implementation was also initiated. The country team developed standards for IDNs and reviewed and released generic TLD and ccTLDs for Mongolian. The local language software developed through the project has been adopted by the government, thus directly reinforcing policy decisions.

5.2.5 Nepal

5.2.5.1 Overview of the Policy Environment

Information Technology Policy 2000 (http://www.unapcict.org/ecohub/resources/nepal-ict-policy) prescribes the political context for undertaking local language computing initiatives in the country. This document defines an aggressive strategy to put the country on the global IT map. The Action plan for implementing the IT policy includes the following activities.

1. Participation of the private sector in infrastructure development
2. Infrastructure development
3. Human resource development
4. Dissemination of IT
5. Promotion of e commerce

Information Technology Policy (2000) identifies the following measure under the Action Plan for dissemination of Information Technology. The following excerpt from the IT Policy 2000 presents the defined measures.

6.4.6   *Content shall be prepared to enhance materials with Nepali materials on the Internet to promote Nepalese arts and culture and to develop rural areas.*

5.2.5.2 Impact of PAN Localization Project's Research on Local Policy

In Nepal, PAN Localization project was executed by MPP, a non-profit library and archive, established to preserve resources in the Nepali language. PAN Localization project team in Nepal was a core member of NLIT, one of the steering committees of HLCIT. During the second phase of the project 2007-2010 however, political instability in the country impacted upon the national plans and ICT sector could not

assume appropriate place in national development. As a result, NLIT also became dysfunctional. However, project team tried to deal with these challenges and effect policy environment of the country.

In the context of expanding policy capacities specifically by improving the knowledge of concerned policy makers in Nepal on local language computing issues, the project team held regular interactions of the country team with the decision makers. During these interactions, the country team got opportunities to present their work and get feedback from policy makers. In September, 2007, the project team in a meeting with policy makers emphasized the development of standards and sound policies for the successful implementation of IT. In February, 2008 standardization of keyboard layouts for Nepal was discussed with the policy makers. In September 2008, computer curriculum in schools was discussed. In March 2009, MPP team and policy makers deliberated on IDNAs for Nepali. In September 2009, MPP team and policy makers deliberated on some proposed issues and MPP also highlighted the progress achieved by PAN Localization project in Nepal.

Efforts on broadening the policy horizon, has been evaluated in terms of the opportunities provided to policy makers for networking and learning about the project's research. NepaLinux developed through the project has been displayed at annual exhibitions like CAN Info Tech organized by the CAN for the last four years which is attended by a number of policy makers. An estimated 3000 copies of CDs/DVDs of NepaLinux both the downloadable and CD/DVD burnt versions have been distributed during these events. In addition for the effective dissemination of the research and to show case its usefulness to the decision makers, NepaLinux was deployed at 10 telecenters (Fulchowki, Dhading, Sindhupalchowk, Myagdi, Kaski, Rasuwa, Dailekh and 8 other places, established partly under the PAN Localization Project (http://panl10n.net), Bhasha Sanchar Project (http://bhashasanchar.org) and other collaborations, and in the process of deployment in another 16 telecenters (Bhaktapur and Butwal and 14 others) under the Rato Bangla Public Private Partnership Network and facilitated by Madan Puraskar Pustakalaya in direct partnership with Nepal Telecommunication Authority.

To evaluate the impact of research work in affecting policy regimes, the most significant contributions of the project was the establishment of Language Technology Kendra LTK which was the collaboration of Department of CS and Engineering, Kathmandu University, Tribhuvan University and MPP). Its mission is "*To carry out research and development activities and actively involved in technology transfer and acquisition in the field of language technologies, to contribute to the advancement of Technologies for the languages of Nepal.*"

In addition, PAN Localization project team in Nepal has also been successful in drawing government attention towards language issues in ICTs. Significant research work was undertaken on Nepali IDNs and language tables and lists of gTLDs and ccTLDs was released.

5.2.6   Pakistan

5.2.6.1 Overview of the Policy Environment

The MoIT is the national focal Ministry and enabling arm of the Government of Pakistan. It looks after all aspects of policy planning and implementation regarding information technology and telecommunications at the national level. National Language Authority (Muqtidra Quami Zaban) is another key organization

established in 1979 by the Government of Pakistan to develop and promote Urdu as national language of Pakistan. Other institutions and organizations significantly contributing to development of local language computing policy are as follows:

1. Ministry of Information Technology and Telecommunications (MoITT)
2. Pakistan Telecommunication Authority (PTA)
3. Universal Service Fund Guarantee Ltd (USF)
4. Pakistan Software Export Board (PSEB)
5. Pakistan Computer Bureau (PCB)
6. Punjab Information Technology Board (PITB), Government of Punjab
7. Sindh Information Technology Board (SITB)
8. IT departments of Provincial Governments
9. Pakistan All Software Houses Association (PASHA)
10. Internet Service Providers Association of Pakistan (ISPAK)
11. Computer Society of Pakistan (CSP)
12. Federation of Pakistan Chambers of Commerce and Industry (FPCCI)
13. Ministry of Education
14. Sindhi Language Authority
15. Seraiki Academy
16. Punjab Language and Culture Center
17. Pashto Academy
18. Balochi Academy

The National IT policy of 2000 is the primary policy statement available on ICT development that frames the political context of local language computing in Pakistan (MoIT, 2000). The focus areas indentified in the policy include *human resource development*, *infrastructure development*, *software industry development*, *hardware industry development*, *Internet, incentives*, *IT promotion & awareness*, *IT usage*, *legislation and regulations*. IT policy also significantly focuses on the development of Urdu and regional language software. The provisions related to localization in this policy are as follows.

 "*Software development is a high growth industry and forms a major segment of the vast IT market and will continue to do so in the future. Integrated efforts to develop software industry with focus on exports (in addition to the local market) would be undertaken. This would include encouragement of local software houses in Governmental projects, local content development, Urdu and regional language software development, promotion of software exports through establishment of International Marketing Network, special bandwidth rates for software exporters, encouraging joint ventures, hiring of international consultants for global business development and fiscal and regulatory incentives for software exporters through State Bank of Pakistan.*" (Software Industry Development)

3.4.7.4 Urdu and Regional Language Software Development

3.3.7.4.1 *Standardized Urdu code plate will be launched and a concerted plan to encourage the development of open source and licensable Urdu software would be undertaken. This will enable plug-ins for popular office and e-mail packages to be made available. This initiative is expected to drive the*

development of other Urdu and Regional software packages for word processing and data base applications.

3.3.7.4.2 The *government will encourage the use of open source operating systems and low cost or free English language Office software for normal operations.*

3.3.7.4.3 The *intent of this initiative is to encourage people to develop skills in working and writing core software for applications and developing tools, which will go beyond the development of the local languages. The application programs for translation, speech to text conversion, databases, ASPs for popular packages will need to be written in currently and newly evolving software.* (IT Policy Strategies)

5.2.6.2 Impact of PAN Localization Project's Research on Local Policy

PAN Localization project in Pakistan has been coordinated by the regional secretariat of the PAN Localization project, at CRULP, NUCES, coordinated by the regional project leader.

In the context of expanding policy capacities specifically by improving the knowledge of concerned policy makers in Pakistan on local language computing issues, the project team held regular meetings with relevant policy making organization to present their research. In 2009, the Punjab Government in Pakistan started Punjab IT labs project to deploy IT labs in 4,286 secondary and higher secondary schools of the province. Thus, CRULP in collaboration with the Punjab IT Labs Project organized a seminar titled "Integrating IT in Education: Language, Curriculum and Training Challenges in Government Schools of Punjab" on 17 December, 2009. The country team presented its experience of end user training on localized ICT in the seminar.

To initiate a public debate on localization in Pakistan, students and teachers were trained on localized technology to generate online content. At the end of the training, a website development competition was organized for the participants. The websites received from the participants were evaluated by a team of experts from the IT industry, academia, government, media and other relevant organizations. With the support of National ICT R&D Fund, Ministry of Information Technology, a prize distribution ceremony was held on January 23, 2010 at the NUCES Lahore campus. This activity created awareness among government decision makers of the work being done under the project and highlighted the important issues to be addressed to promote local language computing policy.

The country team has been involved in developments taking place in IDNs. A workshop was organized by CRULP in year 2008, gathering participants representing the various languages spoken in Pakistan. An initial attempt was made to draft character sets for different languages through this workshop. Two follow up events were organized by Ministry of IT Pakistan to develop IDN ccTLDs for Pakistan.

The other factor specifying the need to expand policy capacity is the organization's capacity to deliver policy relevant recommendations. The project has thus significantly contributed towards improving the capacity of its implementing partner organization to make policy recommendations to their government. In this regard, when the country team was asked to rate their capacity to deliver relevant policy recommendations on a 5-point scale ranging from very limited (1) to very high (5). Pakistan team rated its capacity as very high.

Efforts on broadening the policy horizon, has been evaluated in terms of the opportunities provided to policy makers for networking and learning about the project's research. In this context, the project team has been active in interaction with various professional bodies like Pakistan Software Export Board (PSEB), Pakistan All Software Houses Association (PASHA), Internet Service Providers Association of Pakistan (ISPAK), and Computer Society of Pakistan to realize the significance of the localized technology. In these meetings, project team successfully created awareness about the benefits of localized technology.

The project team also realized the effectiveness media and publicity to influence various stakeholders. A local language website development competition was widely covered by print and electronic media. Spider, popular IT Magazine in Pakistan published articles highlighting progress made by the project in the country. The country team also highlighted its work through its website (http://www.crulp.org/dareecha/). A large number of flyers about the project activities were also distributed. The county team developed books on basic computer skills and localized applications, including Open Office (Word Processor, Graphics Editor), SeaMonkey (Browser, Email Client, and Web Editor) and Psi (Instant Messaging Client). The books created awareness about benefits of local language technology. The team presented its work at different events at national and international levels including the following:

- Sustainable Development Policy Institute (SDPI) Study Group Meeting on Women and ICTS: Exclusion or Empowerment on 13 August, 2009
- Internet Governance Forum (IGF) Workshop on Equality in access to knowledge society, through language and cultural diversity held at Sharm El Sheikh, Egypt on 18 November, 2009
- 2nd GEM Global Exchange, held in Bali, Indonesia from November 24-30, 2009
- Seminar on Integrating IT in Education: Language, Curriculum and Training Challenges in Government Schools of Punjab held on 17 December, 2009 at NUCES, Lahore

To evaluate the impact of research work in affecting policy regimes, the project leader has been actively participating in the consultative process initiated to finalize five year Information Technology Policy 2010 by the Planning Commission of the government of Pakistan. Project leader was also a member of Broadband Stakeholders Group (BSG) formulated by Pakistan Telecommunication Authority (PTA), telecom regulatory body in Pakistan.

PAN Localization project has creditable contribution in Pakistan policy circles regarding the initiation of dot Pakistan registry by Ministry of Information Technology. Another notable contribution of the project was the country team's collaboration with Microsoft for the development of Language Interface Package (LIP) in Urdu for Microsoft Windows Vista and Microsoft Office. In this project, Microsoft Corporation supported the country team to develop interface terminology translation of 300,000 English words into Urdu. These terminology translations have been endorsed by the National language Authority (NLA) of Pakistan and thus the translations developed by the project team have been standardized.

Similarly, the project has also contributed in creating the realization that new IT policy draft must be open for comments from public and other stakeholders. Now, local language computing/ localization have placed in the draft. Advance Technology Program in Human Language Technologies initiated by ICT

R&D Fund is also being considered the outcome of the persistent advocacy made through PAN Localization.

5.2.7 Sri Lanka

5.2.7.1 Overview of the Policy Environment

The ICTA is the apex body involved in ICT policy development. State Language Commission and Department of Official Languages are involved in development of language policy. Sri Lanka Standards Institution (SLSI) is the country's leading standards development organization. The draft National ICT Policy 2006 serves as the main policy statement to understand political context for the project. The draft lays significant emphasis on local language computing as indicated in its excerpt below.

Access to information to be facilitated through the development of ICTs in national languages:

- *Standardization*: Scripts based on the international Unicode / ISO10646 standards, to be used for local language web sites and in the electronic transmission of information.
- *Technical support*: Measures to be taken to ensure that technical support is provided for localization of content.
- *Deployment:* Measures to be taken for the deployment and implementation of standards based Sinhala and Tamil support to relevant sectors and organizations.
- *Content:* Measures to be taken to promote content development in local languages for the Internet and also to facilitate translation of relevant English content to local languages.
- Research to be carried out on the development of content to meet the needs of communities. Community participation in content development in local languages to be encouraged.
- Accuracy, integrity and reliability of information in local languages to be ensured.
- Collection, classification, protection of indigenous knowledge, and making such knowledge available through ICTs in local languages to be promoted.
- Sector based content: Government organizations, universities to engage in development of productivity-enhancing information in local languages relevant to the agricultural sector, fisheries sector; Local language content to be provided on prices of agricultural products, on agro-based industries, and on agri-business development.
- Educational content development in local languages to be implemented. (Sri Lankan Country Report on ICT Localization Policy)

5.2.7.2 Impact of PAN Localization Project's Research on Local Policy

In Sri Lanka, PAN Localization project was executed by UCSC. In the context of expanding policy capacities specifically by improving the knowledge of concerned policy makers in Sri Lanka on local language computing issues, the project team held regular meetings with relevant policy making organization to present their research. The country team has also been actively involved in developments taking place in IDNs. The other factor specifying the need to expand policy capacity is the organization's capacity to deliver policy relevant recommendations. When questioned regarding the organizational

capacity of UCSC to conduct policy relevant research, the project leaders rated it as high (4) on a scale of 1 (very limited)- 5 (very high) points.

Efforts on broadening the policy horizon, has been evaluated in terms of the opportunities provided to policy makers for networking and learning about the project's research. The project leader of PAN Localization project in Sri Lanka was a member IT Sectoral Committee of the Sri Lanka Standards Institution (SLSI) which is a member of ISO and a committee member of ICTPA LLWG, two decision making bodies for IT standardization in the country. He has been successful in effectively using these positions to highlight different local language computing issues to be highlighted and addressed at policy levels. The country team has also been active in creating awareness among community-based organizations CBO's, IT researchers, general public and other stakeholders about the research work being done under the project. The project's research work has also been disseminated the local website (http://ucsc.cmb.ac.lk/ltrl/).

To evaluate the impact of research work in affecting policy regimes, a significant achievement of the project was the development of LIP for Microsoft Windows Vista and Microsoft Office in Sinhala. Through this project, Microsoft supported the country team to develop local language interface for Microsoft Vista operating system and the interface terminology developed has been standardized by ICTPA. In addition the work done by the team on IDNs has contributed to the development of IDN ccTLD registry in Sri Lanka. Further, the country team was actively involved in providing feedback to the government decision makers on important policy issues. The team has and also contributed in constructing policy development process on strong theoretical foundation. The research conducted through the project has been adopted within the government and standardized. E.g. project's research on defining the collation sequence has been incorporated into the SLSI 1134 standard.

National policy necessitating a change to develop government websites in three official languages was influenced by the project work to some extent. The project also contributed to the development of ICTPA LLWG. The working group made significant progress towards localization of key software and government website standards.

6   Gender Mainstreaming in PANL10n

Gender issues exist in nearly every project (Hafkin, 2002) that must be explicitly dealt from the early stages of project. The project monitoring and evaluation system must ensure that gender issues are identified at the outset and appropriate mitigations are strategized within the project phases such that project is implemented efficiently and achieves its stipulated objectives. PAN Localization project aimed to provide equitable access and use of project resources to its project beneficiaries (including both men and women). The gender related project objectives of the project were:

To assess that how far has the project been able to ensure equitable participation of both men and women?

The specific evaluation objectives developed in this context were:

1. To assess the contribution of each country partner institute in mainstreaming gender perspective in the project
2. To assess the gender mainstreaming efforts of country components specifically in technology localization, localized content development and end-user training on localized software
3. To investigate challenges of gender integration in the project phases where gender is least integrated

Nine organizations from six partner countries participated in the subject study. Project partner organization have been strategically sampled in order to ensure adequate coverage of gender mainstreaming evaluation of organizations working in each of the three thematic areas (localized technology, content and training development) of the PAN Localization project.

To evaluate the gender mainstreaming efforts, the PAN Localization project Gender Evaluation Tool has been developed. This tool helped measure the level of gender mainstreaming in each of the project phase as well as the overall project. Specific measures have been incorporated in the tool to evaluate project's gender integration from the respective project's perspectives of localized technology development, content development or end-user training on localized software. The tool is described in detail below.

6.1   PAN Localization Project's Gender Mainstreaming Tool (GMT)

Using the Gender Evaluation tools available in literature (USAID, 2011) as the basis, a gender mainstreaming tool was developed through PAN Localization project for reviewing gender mainstreaming strategies specifically in local language computing projects (tool attached in Appendix E). This tool aims to quantitatively present the level of gender mainstreaming in each of the project phase as well as the overall project. Specific measures have been incorporated in the tool to evaluate project's gender response from three specific localization project related perspectives of; localized technology

development; local language content development; and end-user training on localized software. This tool follows the structural design recommended in the World Bank's *Checklist for the Planning, Design and Implementation of an ICT Project Incorporating Gender Issues.* It further includes the guidelines provided in *Gender Mainstreaming Practices in USAID's ICT related project.* Localization specific measures have been included through the research experiences of the project's regional secretariat staff in executing PAN Localization project.

The GMT for Localization Projects is still an introductory approach. It summarises key steps in applying a gender lens when designing, preparing, implementing and monitoring and evaluating localization initiatives. It further includes guidelines to develop gender responsive project management strategies. Its simplistic approach is a foundation step in developing a comprehensive gender based analytic approach for the development of localization initiatives in the future.

The following sections describe guidelines to follow in order to mainstream gender within each of the project's phases mentioned below.

1. Project Designing
2. Project Preparation
3. Project Implementation
4. Project Monitoring
5. Project Evaluation
6. Project Management

Project Designing

Project designing is the fundamental stage for gender mainstreaming as it determine the eventual access and participation of women in the project. While designing projects, emphasis on gender consideration must be reflected at three levels. Firstly, gender equality must be addressed within the overall project objectives or the terms of reference. This focus can be discretely noted through the project's Vision, Mission documents. If the projects follow the gendered outcome mapping approach, this focus can be evaluated through the analysis of the project's vision, mission, boundary partner and outcome challenge categorization. While framing the project objectives adequate participation of women in the goal setting team is recommended in order to have gender focused project objectives (Hunt, 2004).

Secondly, gender and ICT issues must be explored; specifically those relating to localized technology development, localized content development and training (ICT Tool Kit, World Bank), either qualitatively or quantitatively. These can be observed by meeting the key stakeholders or surveyed through situational analysis, needs assessment studies, baseline surveys, etc. Thirdly the project must design explicit strategies to include both men and women as the target beneficiaries of the project outcomes. In order to achieve this balance, project interventions must be designed to ensure equal participation of men and women (Hunt, 2004).

Project Preparation

Project preparation is defined as an intermediate phase between the project planning and design and its execution. At this stage, project teams must be appropriately educated on gender issues in order to institute gender awareness within the project teams (African Development Bank, 2009). In this regard, gender awareness at an organizational level is considered as a fundamental step for evaluating gender mainstreaming in project. This includes awareness of the project staff on gender issues through workshops and seminars. In addition, projects must have a gender expert in the team to provide guidance from a gender lens during the project phases.

Project Implementation

Practice and literature recommend that women participation in the project is largely detrimental on the congeniality of conditions and environment of the project. The following key steps are therefore recommended within the tool to ensure a conducive environment for women to implement ICT projects.

 "*In the past, it has often been assumed that women benefit automatically from development efforts and that progress towards equality of opportunity and treatment between men and women takes place naturally*"(ILO Evaluation Unit, 2000). Thus the tool recommends development of explicit project strategies focusing on improving women participation in the project. Where it is anticipated that gender issues may hamper women participation, the project activities must be tailored to ensure that gender concerns are appropriately addressed. Similarly, both men and women must be enabled to access and use the ICT resources developed through the project. From a localization project's perspective, technology developed must be of relevance and use to both men as well as women. To facilitate this, adequate partnerships and collaborations are to be developed to enhance the project's outreach and availability of the ICT resources in order to maximize their use by the women.

Project Monitoring

Monitoring is the systematic collection and analysis of information as a project progresses (Hunt, J2004a). It is aimed at improving the efficiency and effectiveness of a project. In the past, project monitoring, emphasized upon the need to collect data on the completion of the project deliverables only. However, recently, development agencies have acknowledged project monitoring as a gender sensitive matter and emphasis is now being laid upon gender segregated data collection (Hunt, 2004) Thus, the gender mainstreaming tool recommends adopting a gender sensitive framework for project monitoring e.g. gendered outcome mapping framework for project planning monitoring and evaluation. Once a gendered approach is used, performance indicators developed must also be gender dis-segregated. This form of data collection would ensure equal emphasis in reporting progress related to the performance outcomes of both genders.

Project Evaluation

Project evaluation following a gender perspective, draws comparisons of the project's strategic plans with its accomplishments and brings up the actual project impact on both gender (Hunt, J2004a). Thus in order to mainstream gendered practices within project evaluation, the tool recommends inclusion of the

following three practices. Firstly, gender analysis or gendered situation analysis of the project must be included in the TORs of the evaluation (Neimanis, 2001). Secondly the project team involved in the project evaluation must be gender balanced (Murison, 2004). Finally, special focus must be paid on presenting and discussing gender based results in project reports (Hunt, 2004). This would require sex disaggregated data collection to effectively reflect upon performance, constraints and challenges faced by both genders in the project (ibid).

Project Management

Gender mainstreaming in project management practices is also fundamental for gender responsive project execution. In this regard, the tool recommends formation of a gender balanced team for executing the project ensuring that equal opportunities are provided for both men and women to work in the project. It is further specified that gender balance is ensured throughout the team hierarchy. It would entail adequate representation of women at the leadership\management positions as well as at the team member levels. Similarly professional development opportunities must also be equally available to both men as well as women.

6.2   Data Collection Methodology

Data regarding the questions asked in the gender evaluation tool was collected through structured questionnaire (attached in Appendix F) filled by the CPI leader. In addition observations and experiences of the project regional secretariat staff and secondary analysis of the project reports was also conducted to report the country specific findings.

Questions asked through this tool were measured against Likert scale (1(low) – 5(high)) and higher ranks were given to the higher gender mainstreaming while lower rank were given to low integration. Each phase of the project was measured independently from other phases for gender integration. Total ranking was calculated by adding the individual phase wise ranks.

After collecting data, empirical analysis was conducted and findings have been synthesized and reported in the section below.

6.3   Country Findings

The following section presents findings on gender mainstreaming in project's country components. Each country section is organized in the following way. Firstly an introduction to the country project is presented. Subsequently, over-all gender mainstreaming rank for the project and individual rank for its phases is discussed. Lastly, specific country case study is presented illustrating benefits of the project outcomes to the women.

6.3.1   Bangladesh

6.3.1.1 Project Overview

Projects in Bangladesh were executed by D.Net and Center for Research in Bangla Language processing at BRAC University, Dhaka, Bangladesh. D.Net focused to make local language information accessible to end-user community including women as well as print-disabled community by carrying out training of content developers and infomediaries to effectively access local language content from the Internet. At the project inception, D.Net conducted a need assessment study to investigate the basic content requirements of rural community for improving their livelihood. Results indicated that major content requirements were those of health, education and agriculture. To make this content available, D.NET established 4 Pallitathya Kendras (telecentres) for delivering content to the rural community. To facilitate rural community in accessing relevant content, D. Net trained infomediaries including both men and women. These infomediaries were tasked to facilitate the rural end-user community coming in a telecenter to access the required content. To establish this relationship effectively, D.Net provided a series of training to capacity build the infomediaries on effective content searching techniques. They were also trained to route any specific content requirements to an online help-desk that they were unable to search themselves. Similarly, content developers were also trained to develop local language content using localized ICT that was requested by the end-user community and unavailable otherwise. Content developers had knowledge of specific content requirement at the rural level and they developed the requested livelihood content. Primarily the content developed through the project was on legal, agriculture and livestock issues. The content was collected from universities and published books by the government and private sectors. Content developers were also trained to use technology for content development like wiki, content management systems and blog.

6.3.1.2 Gender Mainstreaming in the Project

Evaluating the project on the basis of the gender tool the overall gender focus within the project has been noted as "High" and ranked at 4 on a scale of 1-5 as presented in the table below.

Project Phase		Ranking
Project Design	4	High
Project Preparation	5	Very High
Project Implementation	4	High
Monitoring	5	Very High
Evaluation	2	Low
Project Management	4	High
Overall Gender Rating	4	High

Table 6.1: Phase-Wise Assessment of Gender Mainstreaming in D.NET's Project

Table 6.1 illustrates that the gender tool overall assessed the project phases on higher levels of gender mainstreaming within the project. The following section presents the phase wise discussion of salient activities conducted in the project that contributed to this high gender mainstreaming.

In the project's design phase, explicit focus on gender has been planned in the project objectives. Although the project vision and mission statements developed through the gendered outcome mapping framework lacked explicit gender focus, however gender focus is reflected in discrete identification of male and female boundary partners (Male Infomediary, Female Infomediary, Male Content End-Users, Female Content End-Users, Female Print-Disabled and Male Print-Disabled) and their associated outcome challenges.

In addition to the development of gender specific objectives, appropriate women representation has also been ensured in developing the project objectives and execution methodology. Further, gender and ICT issues pertaining to the project were identified during the initial phase of the project, through the needs assessment study on local content requirements of the target end-user community. Lastly the identified target population of project beneficiaries was gender balanced. Equal number of male and female boundary partners was identified through the outcome mapping framework. Thus over-all gender mainstreaming in project's design phase is "High."

Gender Mainstreaming during the Project Preparation phase has been assessed as "Very High (or 5 on a scale of 1-5). This is because D.Net, the organization delivering services in the project was highly gender aware. Firstly the project had appropriately liaison with the regional evaluation consultant on gender issues. The project team notably reports that this involvement *"Helped (them) to incorporate GEM in Outcome Mapping Methodology and provided support to revise M&E plan to collect response from end users differently for male and female."* In addition the project team including the team leads, team members and monitoring and evaluation officers were also provided gender awareness training to help them in executing the project with a gender focus.

During the Project Implementation phase, the project team had made significant efforts to mainstream gender in the project activities. Firstly, the project made explicit strategies to include more women participation in the project. Foremost step taken in this regard was lowering the qualification level of training participants (both content developers and infomediaries) in order to ensure that more women could participate in the training. The project team also carried out advocacy with the local organizations so that they nominate equal number of female participants for receiving the training. When questioned regarding the effectiveness of the strategies to improve women participation in the project, the project team lead attributed it to have helped only *"to some extent."* The project team introduced specific steps to address gender concerns in the project. For example the training of infomediaries and content developer included a specific session on gender issues. Although this session was not directly relevant in content searching, however, it helped introduce gender thinking in the teams. Pre and Post survey of the trainees regarding gender showed that before training majority of the trainees perceived that gender is only related to "*discussion about men and women's rights*" however after the sessions both men and women were more aware that gender also includes discussions about "*social relations of men and women and their rights*".

Although formal partnerships were not developed with key organizations to improve women participation in the project, however as reported advocacy campaigns were conducted for requesting more participation of women in the training teams.

The project provided equitable access to the projects ICT resources to both the male and female participants of the training. Trainees were provided hard printed training kits developed in Bangla, because as reported by the project team, "local language is understandable to all." Similarly, the content developed by the project was also equitably used by both male and female members. As reported by the project team, "*(Based on the content requirements and request of the end-users) it was difficult to get any livelihood content in local language. The content developer had to translate content from English language or had to write content by themselves.*" Owing to content need in the rural community, the project team reported that the end-users accessed the developed content "*to a great extent.*" It was further noted that women end-users mostly accessed "*educational*" type of content, while men had mostly accessed both "*agriculture and health*" related content. Lastly, gender has also been well captured in the project documentation. Based on the review of the project proposal document, project contract, interim project progress reports and final project document, it was analyzed that the country project reports had specific sections addressing gender. Based on the data presented above, the project implementation phase has been reported to have high gender mainstreaming in its activities.

Project monitoring phase had "very high" gender mainstreaming in its project activities. This is because the project had followed gendered outcome mapping framework, a gender sensitive approach for project monitoring.The project team used this framework because "*OMG is an integrated methodology of Outcome Mapping and Gender Evaluation Methodology. By this methodology project member can assess their own activities and find out their project status along with can set new strategy to achieve project objectives. This methodology can also assess project activities through gender lens which helps the team member to incorporate gender issue properly in project implementation.*" Using this framework, the project team had developed gender specific performance indicators by developing different progress markers for the separately identified male and female boundary partners. Based on the development of separate performance indicators for male and female boundary partners, sex-disaggregated data was also collected for project monitoring.

Project Evaluation phase has been ranked at 2 on a scale of 1-5 in mainstreaming gender in the project activities. A fundamental reason for the low score is because gender analysis was not included in the TORs of the project evaluation documents. It was later in the project execution phases that gender focus was included in the project execution. Even though gender analysis was not made a part of the evaluation objectives, however, the project team had a focus not only to evaluate the final gender based project outcomes but also to learn about the capacity building of male and female beneficiaries separately. However, based on collection of gender disaggregated performance indicators being collected in the monitoring phase, sex-disaggregated data was also collected and analyzed for evaluation. The project evaluation team although included adequate participation of women, however lacked any male participation in the evaluation study.

"High" gender mainstreaming has been attained within the project management practices of D.Net. Firstly, specific budget was kept for gender programming within the project activities. Using this allocation, an M&E trained resource could be assigned on the project to develop a sound gender based M&E plan and to design and execute its operational details.

In addition, the, project team composition has also been largely gender balanced with specifically including women at the management positions as well as the team member levels. Fatema Begum Labony, a woman has been leading the M&E of the project, while another woman has been holding a key research position within the project's training execution team. In addition the project has been providing equal opportunities to both men as well as women to benefit from the professional development opportunities available through the regional conferences and training organized by the PAN Localization project's regional secretariat.

A Story of Change: Mobile Lady in Bangladesh

Figure 6.1: Mobile Infomediary helping the villagers

For ensuring effective participation of women and efficient utilization of the project outcomes by the women, D.Net took innovative steps towards attaining these challenging goals. They conceived the idea of Infomediary. It was the model in which rural community especially women were trained to operate technology and use it as per the requirements of community members. Through this facility common women would easily get any information related to their local need regarding work and life. This model made a significant contribution in giving information to the women framers and poor women who could not access this information because of literacy, in availability of technology and no training about its use. By applying this model, it was observed that women awareness was significantly increased. Through the project they had increased knowledge about their rights. They were realizing their legal rights and understanding issues in violence against women. They were acquiring in depth knowledge on how they could make progress and improve their livelihood opportunities.

Based on the analysis above it can be concluded that D.Net's had significantly worked to mainstream gender in almost all of its project phases, beginning from the project conceptualization or designing till project evaluation. The story of change given below presents a case on the successful strategies of D.Net in empowering women to support other women in their community by providing relevant training on content dissemination and development.

6.3.2   Bhutan

6.3.2.1 Project Overview

Bhutan country component of PAN Localization project was executed by the Promotion and Research Division, in the DIT, Govt. of Bhutan. The country component focused its research on development of local language software, training of various end-users on the localized technology developed and some local language content development.

The vision of the department is to create and enable environment for all the Bhutanese citizens to afford and use ICT and harness the benefits of the ICT to enhance their living standard. In cognizant with its vision, the Bhutan country component released the first localized, Dzongkha Linux which was released on 2nd June, 2006. It was a Debian based Linux operating system with localization of Gnome and OpenOffice.org suite. In the second phase most of the work done on Dzongkha Linux has been enhancement of the localization of existing Open Source Softwares. The work has been done in terminology translation of Gnome Interface, OpenOffice.org, FireFox, ThunderBird, Debian Installer and CD burning applications, etc. This has been compiled and released as Dzongkha Linux 3.0.

 In Bhutan, training on installation and operation of Dzongkha Linux was conducted by the DIT, MoIC from 19 to 27 November, 2007. 16 participants of the training represented private IT institutes, education sector, monk body, judiciary and other government organizations. The training program followed the train-the-trainer format. After receiving the training, the participants were expected to work as trainers of Dzongkha Linux in their respective organizations. DIT accepted the responsibility to develop and published the training material and further to mature Dzongkha Linux. The department authorized three IT training institutes representing southern, eastern and central/western parts of the country to conduct training on Dzongkha Linux.

6.3.2.2 Gender mainstreaming in the Project

DIT's project activities were assessed as "high", ranked at 4 through GMT. Phase wise results are shown in table 6.2.

Project Phase	Ranking	
Project Design	4	High
Project Preparation	5	Very high
Project Implementation	4	High
Monitoring	5	Very high
Evaluation	4	High
Project Management	5	Very high
Overall Gender Rating	4	High

Table 6.2: Phase-Wise Assessment of Gender Mainstreaming in DIT's Project

In project designing, DIT was assessed as "high" in mainstreaming gender within the project activities. This was because they included gender equality objectives in the project vision and mission developed through the outcome mapping framework. Further, women participation was included in setting these

objectives. Further, gender and ICT issues related to the end-user training were also identified during project designing. Among the salient issues identified were *High illiteracy level of women in the developing countries; Problems of availability and affordability of software and user support; fewer numbers of women in the formal sector (civil service, private agencies* etc.

As reported, *"Women participation is very low when compared to (men) their counterparts. In every agency, no. of male employees exceeds female employees in Bhutan right now. So almost every time, there were always fewer female participants than men, although opportunities for participation were given equally to both of them. The major reason can be attributed to the presence of gender difference in enrolment of boys and girls: approximately 60% boys compared to 40% girls which is more apparent in remote communities. There, boys are more likely to be sent to schools to study than girls as they stay back to help their parents with the usual agricultural and domestic chores. Overall literacy rate for Bhutan at 59.5 % can also be a major contributing factor. There is also the problem that women are less interested in learning any technology than men."*

Faced with the above challenges, the project team developed measures to ensure a gender balanced target population. As reported these measures included, *reduction in the price of training fee being charged and requested to get equal number of nominations of women and men for the training.* Gender and ICT issues pertaining to localized technology development were not identified. This was because as reported by the team "*While developing Dzongkha Linux software, the team wasn't aware of the presence of different needs of different people based on gender.*"

Gender mainstreaming in project preparation phase has been assessed as very high, ranked at 5 on a scale from 1-5. This was because the project team was holding discussions and consultations with the regional evaluation staff from the onset of the project. As reported this helped the team "*During the selection of trainees, women were given more opportunity to take part in the training. For e.g. while selecting trainees, if the no. of male applicants exceeded the no. of women applicants (which was always the case), we included all the female applicants*'. In addition two members of the project team had received gender based trainings before the project initiation.

Gender mainstreaming has been assessed as "high" during the project implementation phase. As reported by the project team "The end-users that we have been training, in their field of work, again, there are fewer women working in that field. Only in the field of Education and Health, the ratio of women is more than the men. In all other areas the ratio of men working is more than the women are which can be attributed to the same factors as above. Men are more interested in learning and using new technology." Thus the project team employed specific strategies to address these gender concerns. As reported by employing these strategies, the women participation in the training was increased "to some extent."

In addition the project team had developed relevant measures to ensure adequate access and use of the available ICT resources developed through the project. Project team had also established partnerships with other relevant organization e.g. DDC and Sherubtse college to enhance outreach and access of the project resources. Meetings with the key stakeholders in delivering training were held in order to ensure a gender balanced target population. Lastly, project reports also included adequate coverage of gender issues while reporting the project progress.

Project Monitoring phase has been assessed as "very high" in mainstreaming gender focus within the project. This was because the project team was using the gender sensitive monitoring framework, gendered outcome mapping to monitor the project. Sex disaggregated data was being specifically collected and monitored to give special attention to the women performance. These efforts contributed to attain maximum level of gender integration.

For the project evaluation, DIT has been assessed as "high" in mainstreaming gender within its activities. This was because, gender sensitive framework was being followed for project monitoring thus, sex disaggregated information for was also available for analyzing the performance of both the genders separately. Further, women personal was also included in evaluation team for understanding gender issues.

According to the GMT, the project management had "very high" gender mainstreaming in its activities. This was because the project had adequately developed a gender balanced team structure, providing opportunities for leadership and management to women across the team hierarchy. Also professional development training opportunities have been equally imparted to both men as well as women in the project team. And lastly, dedicated budget allocation has been made for implementing gender programming within the project activities.

6.3.3   Cambodia

6.3.3.1 Project Overview

PAN Localization project in Cambodia aimed to empower women through PAN Localization project. To attain this challenging objective, gender mainstreaming approach was followed by the project team to provide equal opportunities to women for the effective utilization of technology. The project team adequately incorporated the gender perspective in each phase of the project. Their effort to mainstream gender has been evaluated through the GMT. The overall gender focus within the project has been observed as "High" by being ranked at 4 on a scale of 1-5. Phase-wise ranking of the project for gender mainstreaming is presented in the table 6.3.

Project Phase	Ranking	
Project Design	2	Low
Project Preparation	5	Very High
Project Implementation	4	High
Monitoring	5	Very High
Evaluation	4	High
Project Management	3	Satisfactory
Overall Gender Rating	4	High

Table 6.3: Showing Phase-Wise Gender Mainstreaming Results in PLC's Project based on the tool

Table 6.3 depicts the overall contribution of PLC to mainstream gender in the project. Phase wise gender mainstreaming results are discussed to locate the grey areas where gender was not integrated properly. In addition, the result of each section has been discussed below in detail.

In designing the project, the project vision and mission statements developed through the gendered outcome mapping framework lacked explicit gender focus. The general project objectives only sparsely refer to inclusion of gender concerns in the project. The project team involved in setting these objectives comprised of an equal number of women. No specific strategies were developed to ensure gender balance target population prior to the project. However, it was aimed to provide equal chance of participation to both men and women. Regarding identification of gender an ICT issues pertaining to technology development, the team had reported that "*the technology is gender neutral.*" In these activities, majorly gender was not incorporated by the project team and eventually GMT appraised this phase having "*low*" gender equality.

During the Project Preparation, PLC ensured that the project team was adequately gender aware. In this context, gendered outcome mapping training was attended by personal of PLC to understand gender perspective. In addition, gender expert was also included in the project to address ICT and gender related issues. These factors contributed in achieving optimum level (5) of gender mainstreaming on GMT scale in its project preparation phase.

In the implementation phase of the project, PLC employed specific strategies in solving gender and ICT related issues. For improving women participation in technology development, as reported, the project team took measures to "Promote and encourage female in software technology development." However as reported the major impediment in incorporating gender concerns in the project has been "*Lack of gender awareness in the team and overall in the region.*" PLC provided equal opportunity to both men and women in terms of availability of the project resources. Further, project team engaged the *ministry of interior and the govt. officials in the national assembly*, the key stake holders' of the technology, in addressing gender and ICT related issues. In addition gender has been adequately presented in the project documentation. By doing these challenging tasks, implementation phase has been ranked as "high" on GMT scale.

Monitoring phase of the project had "Very High" gender mainstreaming in its activities. This was because projected adopted gendered outcome mapping framework, a gender sensitive approach for project monitoring. By adopting this framework, the project team had developed gender specific performance indicators by developing different progress markers for the separately identified male and female boundary partners.

While the project evaluation phase also had "high" gender mainstreaming. This was due to the fact that the evaluation team included gender based analysis in their TORs. However, the evaluation team was not gender balanced. In reporting the findings of evaluation, gender analysis was specifically presented.

Regarding the project management, the tool assessed PLC on satisfactory level. Major reasons for this ranking include the fact that although the women were in a leadership role in the project team however, professional development opportunities were provided more to the male staff than the female staff. Specific budget allocated was also made to facilitate gender related activities.

On the basis of above analysis, it can be concluded that PLC had made significant contribution in mainstreaming gender to the project activities in spite of the project designing. The reason of not integrating gender appropriately in this phase has already discussed.

6.3.4   Nepal

6.3.4.1 Project Overview

ENRD and MPP were the key organizations constituting the PAN Localization project in Nepal. The aim of ENRD was to explore research issues in the effective utilization of the ICT for the rural population in order to make their life easier and more enjoyable. For the attainment of this prosperity, ENRD involved women, students, youth and other groups of rural areas in localized ICTs training. The training was executed by following the top to bottom approach. Firstly, telecenter operators and teachers were trained to use NepaLinux and other localized software so that they could successfully provide technical support and further training to the local community. In second step, every trained teacher/ telecenter operator developed his/her own curriculum and nominated a group of five participants. Each group had representation from target population of women, farmers, students, youth and teachers. Telecenter operator/teacher trained his/her group. They were the local leaders of this program. Ultimately, each local level leader trained his/her own community members for example student leader trained students. ENRD also conducted further trainings on basic computer skills, open office, Net Meeting, Instant messenger and web browser (Mozilla Firefox). MPP has been a part of NLIT, a sub-committee of HLCIT, an apex government body which undertook a number of initiatives aimed at facilitating localization of IT in the country. The country team considerably enhanced NepaLinux (Linux Distribution in Nepali) and also continued its work on terminology translations. The work was done on the development of Nepali spell checker and grammar checker. The team also worked on the development of OCR and NLP for Nepali. The team made progress towards the development of IDN.

6.3.4.2 Gender Mainstreaming in the End-User training project by ENRD

ENRD specifically followed the gender mainstreaming approach to involve women in the project with the spirit of women empowerment through technology. Evaluating the project on the basis of the GMT, the overall gender focus within the project has been observed as "High" by being ranked at 4 on a scale of 1-5. Phase-wise ranking of the project for gender mainstreaming is presented in the table 6.4.

Project Phase	Ranking	
Project Design	3	Satisfactory
Project Preparation	5	Very High
Project Implementation	4	High
Monitoring	5	Very High
Evaluation	2	Low
Project Management	3	Satisfactory
Overall Gender Rating	4	High

Table 6.4: Showing Phase-Wise Gender Mainstreaming Results in ENRD's Project based on the tool

Table 6.4 depicts that project conducted significant activities with paying attention of gender perspective and consequently project attained this high level of gender mainstreaming. These activities are presented in below sections.

In Project Designing phase, special attention has been paid in highlighting gender perspective. The project vision developed through gendered outcome mapping framework had explicit focus on gender. Further, project objectives were being developed by ensuring gender equality in the project. In developing these objectives, women were also equally participated. For making project more effective for women, gender and ICTs related issues were identified. As reported, "*gender and cultural barriers that hamper women to engage themselves fully in the technological world.*" It was also noted that "*the women in target population were not much* literate". Based on these factors, the project selected Mother's group as one of the important boundary partner of the project for the localized ICT training program. As reported, the primary reason behind selecting this group was, "*As mother's group is one of the active community based group running and managed by local people, (thus) PAN Localization training project could (effectively) capitalize (upon) their local network*".

These factors have jointly contributed in the assessment of gender mainstreaming that is satisfactory on the GMT scale.

Gender Mainstreaming during the Project Preparation phase has been assessed as "Very High (an optimum level on lickert scale). This is due to the gender awareness that organization had in delivering services for project. Firstly the project had significantly consulted the regional evaluation expert on gender issues. The project team remarkably reports that this involvement guided them to incorporate GEM in Outcome Mapping Methodology and provided support to improve M&E plan to collect response from end users differently for male and female. In addition the project team had also included a full time gender expert within the project team as well. Based on her recommendation, as reported the project team was able to identify "(i) Gender issues related (to the) training need of team member, (ii) Gender issues related to (project) activities that need to be incorporated in the core programs, (iii) Awareness programs related with gender need to be organized at the local level and (iv) Gender issues related (to) content need to disseminate among the trainees and villagers via telecenter."

Regarding gender awareness of the project staff the team leads, team members and monitoring and evaluation officers had attended gender awareness training for enhancing gender perspective of the project.

While in the implementation phase of the project, the project team had made significant effort to mainstream gender in the activities. The project made evident strategies for incorporating more women participation in the project. These strategies were contributed in involving more women in the project as noticed by the project team. The project provided equitable access to the projects resources to both the male and female participants of the training. Trainees were provided hard printed training kits developed in Nepali, because as reported by the project team, "*It is easy to understand*" and "*It is easy to carry to their home for the study even (when) they are outside of computers.*" Project team specifically focused on different applications in giving training to the different boundary partners. Such as for students project team focus was on giving training in writer, presentation, and e-mail/internet skills and for Teacher focus

was on the spread sheet and presentation, e-mail/internet. While for women team focused on writer and spread sheet. It was ensured that participants both men and women could easily access and utilize the ICT resources equally. These factors have contributed in attaining 'High' level of gender mainstreaming in this phase.

Monitoring phase of the project had "Very High" gender mainstreaming in its activities. This is due to the adaptation of gendered outcome mapping framework, a gender sensitive approach for project monitoring. By adopting this framework, the project team had developed gender specific performance indicators by developing different progress markers for the separately identified male and female boundary partners. Based on the development of separate performance indicators for male and female boundary partners, sex-disaggregated data was also collected for project monitoring.

As compared to the monitoring, evaluation phase could not integrate gender properly as assessed "Low" on the scale. This was because evaluation team did not include gender analysis in its TORs. However gender perspective was adopted in the later stage of the project. Evaluation team also ensured gender equality in its project team.

Regarding project management practices followed by ENRD, gender mainstreaming has been noticed on the "satisfactory" level. Total 3 women out of 10 were part of team delivering end-user training. Within the team, Ms. Ambika Timila was heading the finance department. In addition, the project tried to provide equal opportunity to both men as well as women to benefit from the professional development opportunities available through the regional conferences and training organized by the PAN Localization project's regional secretariat.

Based on the analysis above it can be concluded that ENRD had significantly worked to mainstream gender in almost all of its project phases, beginning from the project conceptualization or designing till project evaluation. The story of change given below presents a case on the successful strategies of ENRD in empowering women to support other women in their community by providing relevant training on content dissemination and development.

A Story of Change

The emphasis of ENRD was to motivate and mobilize women for the participation in PAN project and effective ICTs utilization as women participation in developments was not very significant previously. In this context, they specifically developed methodologies for increasing women participation in the project. To do so, they decided to involve women social activist in the project who were working at micro level for the development of their community.

Mrs. Maya Gurung was one of them, who were working for the progress of her village by involving the villagers in this process. She had completed her primary school and was now engaged in small business. ENRD team involved her in the localization computing training program. After getting this training, she was actively involved in imparting computer education to her community. As she said, "*We hadn't listened anything about computer till today, but we are very happy to get a golden chance to learn about computer. We felt that learning isn't hindered by age. It is very difficult at the beginning but easy afterwards. Now we*

---

can use computer to send news and messages to our relatives and we can know world's news and massages and type many documents with the help of computer. We can use E-mail and Internet for knowing news of different countries it can be received within a minute." The way project integrated women like Gurung, it gave en extra mile to the women participation and project effective utilization at micro level. Now she is very actively involved in giving training to her village members for computer and internet use.

Figure 6.2: Mrs. Maya Gurung imparting computer education

Similarly, the project involved school teachers in this training program; Sarmila was one of them who were trained by the project on basic computer and internet usage. She was working as a primary school teacher in her community and showed great enthusiasm for getting computer education. After that training, she was much motivated to teach others about this innovation and wished to empower her community with contemporary ICTs. She stated, " *Before the establishment of the telecenter in our village, computer was something that we did not know about at all. But then gradually, we discovered that it's not an extraordinary object and hence there was nothing to fear about at all.*"

6.3.4.3 Gender Mainstreaming in the Localization project by MPP

MPP have been actively involved in following the gender mainstreaming approach-a novel for them. However, evaluation of the project on GMT depicts that much effort has been needed to empower women as gender mainstreaming assessed "Satisfactory" on the scale. Phase-wise ranking of the project for gender mainstreaming is presented in the table 6.5. Table 6.5 illustrates the overall performance of MPP to mainstream gender in the project. Phase wise gender mainstreaming results are presented to identify the spaces where objective is attained and where is not. Each section has been discussed below in detail.

Project Phase	Ranking	
Project Design	2	Low
Project Preparation	5	Very High
Project Implementation	4	High
Monitoring	5	Very High
Evaluation	4	High
Project Management	2	Low
Overall Gender Rating	3	Satisfactory

Table 6.5: Showing Phase-Wise Gender Mainstreaming Results in MPP's Project based on the tool

During the designing of the project, explicit focus on gender has been planned in the project objectives, although the project vision and mission statements developed through the gendered outcome mapping framework lacked explicit gender focus. Further, project did not involve any women in setting these objectives. No specific strategies were developed to ensure gender balanced target population prior to the project. However, it was aimed to provide equal chance of participation to both men and women. In these activities, majorly gender mainstreaming was not done by the project team and ultimately GMT assessed this phase having "low" gender equality.

In Project Preparation, MPP personal attended gender awareness training. In addition, Project Manager-Bal Krishna also participated in gender outcome training workshop at Regional Secretariat of PAN Localization Project. Gender equality and women empowerment was advocated by him after getting this training in his organization.

While in the implementation phase of the project, MPP developed specific strategies to address gender and ICT related issues. Project team also developed printed and soft kits of the training program for making training more effective for the participants. These materials were developed in both local and English langue. The content and technology that created during this project was gender neutral. MPP tried to make the availability of the project resources possible for both men and women and they succeeded in attaining this objective. Project also pointed out some gender and ICT issues that were lack of graphic interfaces for women, lack of automatic translation software, unavailability of computers and internet to the women, and dominance of English language used in technology that hinders most of the women. These factors were addressed during this project. Further, consultation was also done with other key stake holders in addressing gender and ICT related issues. By doing these challenging tasks, implementation phase has been ranked as "high" on GMT scale.

For the project monitoring and evaluation, MPP adopted gender outcome mapping framework. Project team had developed gender based performance indicators to observe different progress markers that were developed separately to identify male and female performance. Sex-disaggregated data was also collected on the basis of the separate developed performance indicators for male and female. Optimum level of gender mainstreaming was attained on GMT scale because of doing those activities. While evaluation was also had "high" gender mainstreaming. It was due to the evaluation team included gender based analysis in their TORs. Evaluation was also done on the basis of gender but the evaluation was not gender included women in this process. Management phase of the project also had "low" gender mainstreaming. This is because project did not include women in this phase. However, the budget was allocated for gender activities in the project.

As above analysis shows, it can be synthesized that MPP had made significant contribution to mainstream gender in its project activities despite the project designing and management. The reasons of not incorporating gender fully in this phase have already discussed. This discussion would help them to mainstream gender in its upcoming projects.

6.3.5   Pakistan

6.3.5.1 Project Overview

CRULP executed the Pakistan country component of PAN localization project. Regarding localized technology development, CRULP had made significant efforts to raise technological support for Urdu in computers. The team worked on the development of IDN and Locale. The team also did significant research in terminology translation.

 The country team also focused on promotion of digital literacy using the localized technology. End-users including students and teachers from rural areas were trained for emailing, accessing internet and local content publishing through websites. Training material was also developed in Urdu. For content access-Browser (SeaMonkey Navigator) was localized. For communication - email client (SeaMonkey Mail and News Group) and instant messenger (Psi) were localized. For Content generation – word processor (OpenOffice.org Writer), graphic editor (OpenOffice.org Draw) and webpage developer (SeaMonkey Composer) were localized. Localization was carried out by collaborated efforts of technical (Software engineers) and non technical (linguists) resources. A training book for each localized software application was developed. To gauge the competency of target end users a competency scale was also developed for each application.

Ten rural schools were selected for training. The medium of instruction in these government schools was Urdu. The schools included 5 boys' schools and 5 girls' schools. Separate training sessions were conducted for teachers and students. Three training phases were scheduled. In first phase, teachers and students were trained for basic computer skills and web browsing. In second phase, they were trained for email and instant messaging. In third phase, they were trained for word processing, graphic editing and web page development. The country team also explored the ways in which the gender dimensions could be integrated into the activities of the program.

6.3.5.2 Gender Mainstreaming in the Project

Gender mainstreaming has	Ranking	
Project Design	4	High
Project Preparation	5	Very High
Project Implementation	4	High
Monitoring	5	Very High
Evaluation	4	High
Project Management	4	High
Overall Gender Rating	4	High

Table 6.6: Showing Phase-Wise Gender Mainstreaming Results in CLE's Project based on the tool

Gender mainstreaming during the project design phase has been assessed as "High" as shown in table above. This is because CRULP explicitly focused on gender during setting the project objectives. The project vision and mission statements developed through the gendered outcome mapping framework also reflected explicit focus on gender. The vision of the project included "*Localized software are reducing the intensity of technophobia especially among females, so they are no more self excluded from ICT and is*

thereby decreasing socio-cultural barriers. By utilizing localized software, they are not only becoming a part of networked world but also generating content in local language which is ultimately helping to decrease the level of digital divide". The mission of the project also has a clause "*the program will contribute to reduce gender gaps by providing equal opportunities for both genders*".

CRULP throughout PAN localization project paid particular focus on the participation of women and interestingly, according to gender tool, this lead CRULP's to have gender imbalance in the project team (by having more women than men) during setting objectives and execution methodology. This is one of the reasons to not attaining the optimum gender mainstreaming during project designing phase.

Moreover, CRULP also focused on identifying gender and ICT issues prior to implementation of the project. Gender and ICTs issues were identified through situational analysis of the target end-user community. The team also identified few gender concerns themselves by being the part of the same culture. Although, the project did not identify gender issues pertaining to access and use of software to be localized and deployed, yet particular attention was paid during the development of training material. After identification of gender and ICT issues, the project proactively designed intervention to specifically include both men and women in the project. Thus, the identified target population of project beneficiaries was gender balanced. By using outcome mapping framework, the project identified equal number of male and female boundary partners and also their associated progress markers. The boundary partners included female students, male students, female teachers and male teachers.

Project preparation phase has been assessed as "Very High" or 5 on a 5-point scale (1-5). This is because CRULP, the organization delivering services in the project was highly gender aware. Firstly the project had appropriately liaison with the regional evaluation consultant on gender issues. During the initial phase of the project, CRULP organized a seminar on gendered outcome mapping in which a special session on gender was conducted for team members including training and evaluation staff. The purpose of this seminar was to raise gender awareness among staff and also to assist them in executing the project with a gender focus. During the project, the evaluation officer also attended GEM Thematic Adaptation Workshop for Localization Initiatives, Jan 17-19, 2009 Vientiane, Lao PDR.

In project's Implementation phase, CRULP achieved "High" (or 4 on 1-5 scale) level of gender mainstreaming. This high level gender mainstreaming exhibited that the project team had made significant efforts in addressing gender equality. The project team employed explicit strategies to address gender concerns and specifically include more women participation in the project. As reported, foremost steps taken in this regard were: " *i) selected equal number of boys and girls schools for the training, ii) formation of gender balanced project team in the context of local norms, the female trainers conducted the training in girls' schools and the male trainers conducted the training in boys' schools in parallel, iii) Use of positive words to avoid the negative stigmas attached to them e.g. instead of chatting the project use the word of instant messaging.*" When questioned regarding the effectiveness of the strategies to improve women participation in the project, the project team lead attributed it to have helped "*to a great extent*" (or 4 on a 1-5 scale)."

Among one of the fundamental reasons in achieving high level of gender mainstreaming during implementation phase is the provision of equitable access and use of project ICT resources to both men

and women participants of the training. For instance, the training program was implemented at ten schools in rural areas of Pakistan, five of these were boys' schools and five were girls' schools. The fundamental reason of selecting equal number of boys and girls schools was to ensure that both male and female students had equal opportunity to access and use localized software. The trainees both men and women were also provided hard printed training kits developed in Urdu (local language), because as reported by the project team, "Urdu is effortlessly understandable as compared to English for rural dwellers specifically where project was executed" . The project team also reported that one of the reasons of providing hard printed manuals to trainees because "this was the standard medium in schools". To ensure the gender equality in the project, CRULP also developed formal partnerships with key organizations specifically with Government officials DCOs.

Gender mainstreaming during project monitoring phase has been assessed as "very high". One of the main reasons in attaining this very high level is because of using OMg Framework. As reported the fundamental purpose of employing this framework was because *"OMg can assess project activities through gender lens which helps team member to incorporate gender issue properly in project implementation."* Using this framework, the project team had developed gender specific performance indicators by developing different progress markers for the separately identified male and female boundary partners. Based on the development of separate performance indicators for male and female boundary partners, sex-dis aggregated data was also collected for project monitoring. Based on the review of the project proposal document, project contract, interim project progress reports and final project document, it was analyzed that the country project reports had specific sections addressing gender.

In project's Evaluation phase, CRULP has been attained "High" level gender mainstreaming. This is because gender analysis was included in the terms of reference of the project evaluation documents. Owing to this, project evaluation team focused to evaluate gender based project outcomes. CRULP also collected and analyzed sex based disaggregated data. Project interim reports, progress reports and research publications included specific sections highlighting gender issues. The fundamental reason was not achieving "very high" integration level is because the evaluation team was not gender-balanced. The evaluation team based on only female evaluators and there was not a single male evaluator in the project.

Gender mainstreaming for project management has been assessed as "High". CRULP allocated specific budget for gender programming within the project activities. CRULP also provided equal opportunities to both male and female team members for professional development available through the regional conferences and training organized by the PAN Localization project's regional secretariat. For instance, Ahmed Muza, a male, participated in OCR training which was held at NECTEC, Thailand from 2nd June 2007 till 28th July, 2007. Three males and four female team members attended "Regional Conference on Localized ICT Development & Dissemination across Asia, 11th -16th January, 2009. Sana Shams and Mudasir Mustafa, two females, participated in GEM Thematic Adaptation Workshop for Localization Initiatives organized at Vientiane; Lao PDR dated Jan 17-19, 2009.

CRULP also paid particular attention for the composition of gender balanced team. Yet, the project team has been attained only 4 on a 1-5 scale. The data shows an interesting finding for not achieving optimum level on 1-5 scale that there were more females involved in Pan Localization project as compared to male team members. This is also worth mentioning that the leading positions were hold by female members as

Sana Shams, a women has been leading the M&E of the project, while another women Huda sarfraz, has been holding a "team lead" position within the project's training execution team. Based on the analysis above it can be concluded that CRULP's had significantly worked to mainstream gender in almost all of its project phases, beginning from the project conceptualization or designing till project evaluation.

A Story of Change

CRULP paid special attention to the women participation in the project; they adopted OMg for integrating women participation in their project because of current practices in which women and girls always being neglected. However, if they participate in any program then they have to ensure that their participation must not cross the cultural constraints especially in rural areas. When project team went to the rural areas, they had that scenario in their minds. In giving trainings to the teachers and students for internet usage and website development, project team headed by Huda Sarfraz – a 29 years old young lady treated their training group on gender basis and divide into four groups that were boy students, girl students, men teachers and women teachers.

Project team trained those school children and teachers belonged from the rural Punjab to use the internet so that they could eventually create their own content. Eventually, all of them created content with their new skills, students and teachers in rural villages created 57 new, locally-relevant school and community web sites, which they presented in a competition held by Dareecha (new window). The judging panel, comprising government officials, academia and ICT experts noticed the strong presence of women and girls among the winners. This was a due to adaptation of gender based GEM was used by the project team of to complement other planning methods for the project, was helped them get through to a segment of the population that other more traditional planning methods may not have achieved: women and girls.

Figure 6.3: Rural female students getting computer education

To meet up with the diverse needs of both sexes meant that girls as well could learn about computers—and the young girls were not letting this opportunity go. "Women teachers were quite insistent in getting the girls involved," recounts Sarfraz "and there was more than one incident where girls were in tears because they wanted to join sessions that were already full". The quotas were the same for boys and girls which meant that there were about fourteen students per session working to a computer. Because the girls were so persistent, a second, afternoon session was set up for a total of 30 girls. "We didn't know what to expect but the girls were very insistent. We were also surprised that the parents were so willing to let

them come to classes during the summer and at irregular hours," explains Sarfraz. She also recalls the story of two girls, the daughters of the head master at a boy's school who felt the training was so important that he took it upon himself to bring his daughters to the boys' training. "I asked our trainers about how the boys reacted to this and was told that the head of the school was a much respected teacher, so nobody said anything to the girls and things went on as usual". The fact that the girls were accepted without question in the boy's classes represents a significant shift in mentality by the community, and for the father himself. "For me, it was the best part of the project," beams Huda. Project aimed to facilitate all the participants but in some cases where special arrangements were made to train them. This strategy became significant when one of them, a women teacher later attended one of the sessions and the awards ceremony in Lahore. It was something that was real a milestone.

In a country where rural girls' schools are sometimes burnt down, least opportunities provided for the education, scarified them for boys higher education and community not favor to educate women. The experiences of the project team were astonishing as they were never rejected by the villagers.

Burning schools is one extreme, more common is a generally negative sentiment towards women gaining access to technology and new communication channels, because there is fear that the exposure to new ideas and people will have a negative impact on girls. The desire to participate by so many girls in the program, and their strong presence at the awards ceremonies (both as attendees and award recipients) sent out a clear message: we are willing and capable of taking part in the information society. All of these surprising milestones were not came out accidently, it was due to the adaptation of gender mainstreaming parameters reflects in GMT for project planning, execution and evaluation.

7   Conclusions and Recommendations

The purpose of this evaluation study has been to assess the effectiveness of PAN Localization project in building research capacity in the project teams, explore how localized ICTs have facilitated the community to learn and use ICTs, access projects' contribution in influencing policies in the partner countries and to assess the project's efforts in mainstreaming gender.

This section is divided into four sub-sections. Section 7.1 presents a recommended sequence for research capacity building principles specifically in localization projects based on the project's experiences. Section 7.2 presents that language is a significant factor in improving digital literacy and diminishing the gender digital divide. It further explains that the education level of the community is another factor contributing in the success of learning to use localized ICTs. Section 7.3 highlights the key findings regarding effectively influencing the policy making in localization. The last section, Section 7.4 recommends key actions required for mainstreaming gender in localization initiatives.

Sequence of RCB Principles for Localization Projects to Achieve Maximal Impact

Though the six principles of RCB holistically address the challenge of RCB in localization, however appropriate sequencing within the six principles must be done in order to foster maximal impact (Potter & Brough, 2004). Based on the project experience RCB for localization must follow a developmental cycle within the defined capacity building principles.

Initially *Skill Building* and *Infrastructure development* must form the focus of RCB interventions for localization. Different countries require different research skills and infrastructure needs owing to the existing competencies. Thus a participatory need analysis should be performed to ensure skill development is based on national priorities and capacity. Based on the identified RCB needs, appropriate mentorship structure and organizational resources may be planned to ensure development of the research base. Secondly, localization RCB can be built to carry out *close to practice research*, of direct benefit in practice, after development of technology the basic and intermediate levels of localization research as a pre-requisite, only then can the research be conducted that harness solutions that can be readily used by the benefitting populations. At the same time, capacity building initiatives must target to form *linkages and partnership* with relevant academic, policy making, regional standardization bodies, and public and private sector bodies. This would follow skill development at level 1 as synergetic and mutually benefitting collaborations can only be developed if the local teams are able to contribute back to the knowledge network. Finally, *research dissemination and sustainability* must be targeted as these RCB dimensions provide research maturity to publish in technical forums, and compete for funding.

Language of ICTs to improve digital Literacy

In general, it has been observed that training on localized ICTs significantly enhanced the capacity of end-users to use ICTs. The findings revealed participants who had never used a computer before this

training were not only using the computers for routine tasks but also providing trainings to others. This is worth mentioning that this ability and confidence to train others was developed soon after receiving the short training on ICT skills through PAN Localization project. E.g. in Nepal, the team adopted the "train the trainers" approach for training program. The project team conducted a 10 days training using localized ICTs to five participants. Each trainee after 48 days of this training program prepared his/her own training content and further trained five more rural community people (including farmers, students, teachers, women etc.) in their village. Then, each participant trained by them, further provided training to five more persons and prepared his/her own training material to impart trainings to them. The monitoring and evaluation officer of the Nepal team reported, "*End-users are "**highly** (ranked 4 on 1-5 scale) convenient in learning a software that has been localized versus the one available in a foreign language and localizing the software has **significantly** (ranked 5 on 1-5 scale) helped end-users to use software*". Similar training approach was adopted and similar findings were observed in Cambodia. Through project experiences in Bangladesh, it was noted that the local language content created by team was "*highly accessed*" by rural community as ranked 4 on 1-5 point scale. Their skills to develop local content were also enhanced as most of the content developers were not familiar with Joomla, blogging and wiki before training. After training, almost every content developers developed his/her own blog and in local language such as: i) http://prothom-aloblog.com/, ii) http://www.somewhereinblog.net/blog/tanzin2008/28783455, iii) http://www.prothomalo.com/ and iv) somewhereinblog.net/blog/kibriyarafiblog. Most of the respondents reported that localized ICTs facilitated them to develop content easily in local language and they also stated, "*by using Joomla, Wiki and Blog in Bangla language, content can be develop swiftly*". Similarly in Pakistan, majority of the trainees were either not familiar with computer or had never touched or used computer before the ICT training introduced by the PAN Localization project.   Yet, within a period of eight months, the students and teachers were able to develop websites using computers. 56 websites were developed by the participants of which one website was developed in English language and 55 websites were developed in Urdu language. Thus it can be firmly stated that *the fundamental factor that enhanced the skills and confidence of trainees to use ICTs is the provision of ICTs in the "local language".*

Association of Education and Language of ICTs in Learning ICTs

Learning to use ICTs requires the competency to read and write English which is lingua franca of ICTs. It is widely recognized that education significantly helps to improve the ability to read and write English language which in turn means that to utilize the benefits from ICTs one must has to be well-educated or highly qualified.   Data was collected from the project trainees to quantitatively measure their reading and writing skill presented in the figure below provide two main deductions. First, local populations, irrespective of their education levels, are more competent in reading and writing local language as compared to English or any secondary language. Second, the level of education affects their competency level to read and write English. This is because survey respondents having higher education levels were more capable to read and write English as compared to the respondents who had less formal education.
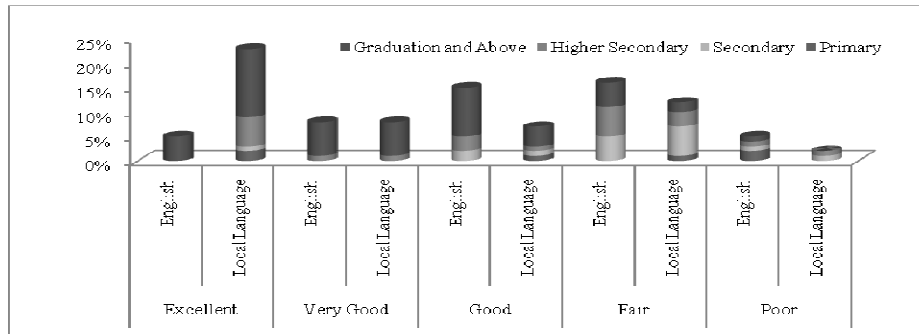
Figure 7.1: Education and Language Association for Reading and Writing Skill

Figure 7.1 above shows the relationship between educational levels and self-assessed competency levels in reading and writing English and local language. The figure shows that the respondents who rated them as "excellent" in reading and writing English had "graduation and above" education. Thus, it is clear that it is difficult for the persons who have less education to effectively get the benefits from the ICTs in foreign languages. In addition, the results about to use the computer with English interface or local language interface shows that the respondents who had "graduation and above" education, significantly preferred to use computer with English interfaces, than those who had "secondary" significantly preferred to use computer in local language interface.

Thus, the findings reveal that ***it is difficult to learn ICTs in English specifically for those who have lower education levels***. They may learn but cannot effectively get the benefits from ICTs as according to Benson (2004), the program based on the instructions through a language that learners do not speak may succeed teaching student to decode words, yet it can take years before they discover meaning in what they are reading.

Gender and Adoption of Localized ICTs

ICTs in English place women in a more disadvantaged situation than men, as for instance; according to the GPI computed for South Asia, only 51 percent of women had the ability to read and write in local language. Thus ICT initiatives will not be effective unless they deliver information that is useful and relevant to the end-users and where the end-users (women and men, girls and boys) have the capacity to act on it (Primo, 2003). Through PAN localization project the ICT literacy training program initiated had an active participation of women in every country.

Provision of computers with local language interfaces and availability of content in local language was the dominant reason for this active participation of the women, This is further strengthened by the project findings presented in Chapter 4 presents that women significantly preferred to use local language for online information exchange as well as for computer interface. The content developed in local language by Bangladesh component was highly accessed by community reported by the Bangladesh team. This is because the content was in the language of the native community. The study reported by Morna and Khan (2000), explains one of the reason of the low attendance at tele-centres in Africa is the language and content that does not speak to the language of the end-users community. ***Thus, availability of localized ICTs and content in local language can significantly help in bridging the gender digital divide.***

Communication for Policy Influence

Ineffective and untimely communication of research has been regarded as a key attribute because of which research work was ignored and could not manage to influence the policy and practice in a study covering over 70, 000 research project in education (Molas et al, 2000). PAN Localization project had foreseen this risk at the time of project planning. Thus as a strategy, each country team was asked to submit a report on "Meeting with policy makers" on six monthly basis along with the other project's research deliverables, which were eventually tied up with the release of funds for the next term. ***Holding frequent interactions with the relevant organizations, help in disseminating project's research in policy circles*** and gain their feedback on policy directive for aligning the research with policy. A short coming in this approach, however, has been that the project did not equip its teams with a communication for planning training at the outset of the project, which could have provided a more concrete and focused approach to the discussions.

Taking the Decision Makers Along

A significant factor attributed in the ignorance of research in policy and practice described by the study conducted by Molas et al (2000), is the "inadequate supply of, and access to, relevant information [by the policy makers]." In this context, the project had organized a regional meeting on "Sustainability of Local Language Computing across Asia" in Bhutan, in Jan, 2007. At this forum, the project brought together senior national policy advisors from more than a score countries to discuss the existing national policies on localization in the regions and to debate the best policy for the future of the region. These policy makers interacted with the country project leaders over three days formally and informally and came up with a set of recommendations. A real achievement of this consultation has been that it gave countries to see what successful models are and how they are working. Through this interaction, they also realized the resources such measures require and the timelines they need. This interaction further helped provide the complete understanding to the research challenges, required solutions and the role public organization could play in fostering sustainable development in the area. ***Thus it is fundamental for the success of the research to keep decision makers known about the research issues, challenges and solution, in order to enable them to stipulate the right decisions***.

Successful Projects Attract Policy Makers' Interest

The success of PAN Localization project has contributed in initiation of further research and demand for local language computing projects nationally as well as at the regional level. A significant impact of the success of PAN Localization project has been in developing the confidence of key stakeholders at both national and international level in the technical ability of the project team and thus their focus to further fund in this area. A concrete example is the grant of localization project to PAN Localization project partners by Microsoft for developing the Language Interface Package (LIP) for Microsoft Windows Vista and Microsoft Office in Pashto (by Afghanistan team), Bangla (by Bangladesh team), Urdu (by Pakistan team) and Sinhala (by Sri Lanka country component) etc. The project has also influenced the governments to develop action plans to give out local language computing projects. There have also been contributions in policy decisions/ decision making process. The project made significant progress toward shifting focus of governments on language issues in ICTs. Two good examples are Pakistan and Nepal. The project has also been successful in introducing changes in policy development process in terms of

inclusion or exclusion of certain organization, criteria, membership and representation. In Pakistan, public participation in policy development process can be considered one of the most significant achievements. Pakistani team was actively involved in developments taking place in IDNs. This work invited public participation through the workshop organized in 2008, gathering participants representing the various languages spoken in Pakistan. *Successful projects pave way for further funding and inviting policy makers' attention to the research issues*.

Is Technology Gender Neutral?

Gender is predominantly present in every project. Project may fail to analyze gender issues because projects are not examined from that aspect. This is because every project is impacted by its socio cultural contexts in which gender is central. Failure to consider gender differentiated impact may have negative impact on the project outcomes. The following table presents a summarized gender mainstreaming view of the case studies presented in the country findings' sections, Chapter 6 above.

Phase	KH	NP –ENRD	NP –MPP	BD	BT	PK
Design	2	4	2	4	4	4
Preparation	0	5	5	5	5	5
Implementation	4	4	4	4	3	4
Monitoring	5	5	5	5	5	5
Evaluation	3	1	4	2	4	4
Project Mgmt.	3	3	2	4	5	4
Overall	3	4	3	4	4	4

Table 7.1: Summary of the Assessment of Gender Mainstreaming in PAN Localization Project's CPIs

The assessment results shown above elucidate higher levels of gender mainstreaming in end-user training projects versus lower level of gender mainstreaming in the localized technology development projects. The end-user training teams included country teams in PK, BT, NP-ENRD and BD-D.Net which presented "high" gender mainstreaming across the projects while the remaining teams where focused more on localized technology development that could only attain "satisfactory" level gender mainstreaming in the activities. As already discussed, the major reason for this result is the attribution of technology as gender neutral by most of the project teams developing localized technology and no stories of change are available to present how the local language application development has been an enriching experience for women involved in the project*. More research thus needs to be concentrated in defining the literature on how gender dimensions could be explicitly considered in the software requirement gathering, design and quality assurance.* When gender was adequately addressed in projects, especially the end-user training projects, various positive impact of ICT on women can be seen, in terms of motivation, increased self esteem and self confidence. Women who learn to operate technology feel empowered and are motivated to champion similar efforts in their community promoting both a multiplier effect as well setting a positive precedent for others.

Explicit Strategies for Gender Mainstreaming must be developed

It has been observed that countries that adopted a systematic framework for gender mainstreaming presented better gender mainstreaming results than teams that did not. In spite of the fact that

implementation of gender responsive framework in projects require training in the area, nevertheless expertise once built can be effectively utilized in other domains as well.

It has been observed that project leadership of the technology development project usually agree to inclusion of a gender balanced team for project execution. However, while going through the staff hiring procedure end-up recruiting male staff for certain technical localization related positions creating an imbalance as in the case of MPP, Nepal. However in such a scenario strategies need to be employed that ensure that women are hired and trained for such technical positions by creating liaisons with university staff or other academic institutions. Equal opportunities for profession development need to be provided to both girls and boys alike in order to ensure gender positive outcomes.

From a logistical perspective, sometimes explicit budget provisions have to be incorporated in the project if the project needs to safeguard women participation in outreach projects. From experience of ICT training program held in rural areas of Pakistan it was found out that sometimes to ensure participation of a female teacher at an out station venue, her family member accompanying her had to be accommodated. This is because of cultural constraints that require a female to be accompanied by a family member in most cases, thus limiting the physical mobility of women. ***Thus gender mainstreaming in ICT localization project is only possible if explicit measures to ensure gender integration within the project are devised***.

References

African Development Bank (2009). *Checklist for gender mainstreaming in the Infrastructure sector*. Retrieved from http://www.afdb.org/fileadmin/uploads/afdb/Documents/Policy-Documents/Checklist%20for%20Gender%20Maintstreaming%20in%20the%20Infrastructure%20Sector.pdf.

Accenture & Markle Foundation. (2001). Creating a development dynamic: Final report of the Digital Opportunity Initiative Retrieved from http://www.opt-init.org/framework/DOI-Final-Report.pdf
Annan, K. (2002). Information and communication development: Information society surmit. P.7

Ansari,S., & Saleem, S.(2009). Pakistan. In Akhtar, S. & Arinto, P. (Eds.), *Digital Review of Asia Pacific*, 2009-10 (p. 294-301). New Delhi: Sage Publications.

Ariunaa, L., & Uyanga, S. (2009). Mongolia. In Akhtar, S. & Arinto, P. (Eds.), *Digital Review of Asia Pacific, 2009*-2010(p. 268-273). New Delhi: Sage Publications.

Arrison, S. (2002). What digital divide? Retrieved from http://news.com.com/2010-1078-858537.html

Auriat, N. (1998). *Social policy and social enquiry: reopening debate*. UNESCO, Oxford: Blackwell Publishers.

Baliamoune-Lutz, M. (2003). An Analysis of the Determinants and Diffusion of ICTs in Developing Countries. *Information Technology for Development* 10(3), 151–169.

Batchelor, S., Norrish, P., Scott, N., & Webb, M. (2003). *Eldis Sustainable ICT case histories*. Retrieved from http://www.sustainableicts.org.

Bayanduuren, D. (2007). *PAN Localization Project Phase II, Mongolian Speech Recognition.* Retrieved from *www.**pan**l10n.net/Presentations/Bhutan/**Phase2**/MUST.pdf*

Benson, C. (2004). *The importance of mother tongue-based schooling for educational quality*. Retrieved from http://unesdoc.unesco.org/images/0014/001466/146632e.pdf

Breen, CM., Jaganyi, JJ., Wilgen, V. BW., & Wyk, V.E. (2004). Research Projects and Capacity Building. *Water SA* 30(4), 429-434.

Bureau of Democracy, Human Rights, and Labor. (2010). *Human Rights Report*. Retrieved from http://www.state.gov/g/drl/rls/hrrpt/2010/sca/154477.htm]

Bureau of East Asian and Pacific Affairs .(2011). Background Note: Laos in U.S department of state. Retrieved from http://www.state.gov/r/pa/ei/bgn/2770.htm

Burnett, G. (2000). Information exchange in virtual communities: a typology. Information Research, 5 (4). Retrieved from http://informationr.net/ir/5-4/paper82

Caplan, N. (1979). The Two-Community Theory and Knowledge Utilization. *American Behavioral Scientist*, 22 (3), 459-470.

Carden, F. (2004). *Issues in assessing the policy influence of research*. UNESCO, Oxford: Blackwell Publishers.

Connolly,P. & York, P. (2002). Capacity-Building efforts for Nonprofit Organization. *OD Practitioner* 34(4), 33-39.

Cooke, Jo. (2005). A Framework to Evaluate Research Capacity Building in Health Care. *BMC Family Practice*, 6(44).

Coomaraswamy, R. (2005). The varied contours of violence against women in South Asia. presented at the *fifth South Asia Regional Ministerial Conference*, Celebrating Beijing Plus 10. Islamabad, Pakistan.

Crewe, E., & Young, J. (2002). Bridging research and policy: context, evidence and links. *ODI Working Paper 173*. Overseas Development Institute: London.

Cueto, S., Guerrero, G., León, J., Seguin, E. & Muñoz, I. (2009). Explaining and overcoming marginalization in education: a focus on ethnic/language minorities in Peru, background paper for EFA Global Monitoring Report 2010.

Darrow and Saxenian. (1998) Article: Victor van Reijswoud, Divine Word University, Papua New Guinea Retrieved from http://www.ejisdc.org/ojs2/index.php/ejisdc/article/view/548/293.

Department for International Development. (2008), Research Strategy 2008-2013. Working Paper Series: Capacity Building. Retrieved from http://www.dfid.gov.uk/r4d/PDF/Outputs/Consultation/ResearchStrategyWorkingPaperfinal_capacity_P1.pdf

Department for International Development (DFID). (2010), *How to note Capacity Building in research. Retrieved from* http://www.ukcds.org.uk/_assets/file/publications/HTN-Capacity-Building.pdf

Donny B.U., & Mudiardjo, R. (2009). Indonesia. In Akhtar, S. & Arinto, P (Eds.), *Digital Review of Asia Pacific*, 2009-2010 (pp. 201-209).New Delhi: Sage Publication.

Dijk, J.V., 2008. The Digital Divide in Europe. The Handbook of Internet Politics, Routledg. London and New York. Vol. 19, pp. 315-326.

Doran, J. (1996). *Rural Transport*. London, Intermediate Technology Publications.

Earl, S., Carden, F., & Smutylo, T. (2001). Outcome Mapping: Building Learning and Reflection into Development Programs. Ottawa: IDRC Canada.

Ethnologue. (2005). *Languages of the World*.   SIL International , 15th Edition.

Global Reach. (2011) Retrieved , from http://www.glreach.com/globstats/index.php3

Garnham, N. (1997). Amartya Sen's capabilities ap-proach to the evaluation of welfare: Its applica-tion to communications. The Public, 4 (4), 25–34

Government of Bangladesh & United Nation. (2005), Millennium Development Goal. Retrieved 2011, from www.searo.who.int/LinkFiles/MDG_Reports_BangladeshMDG.pdf

Gul, S. (2004). Dilemmas of Localisation in Asia. *I4D Online*, 2(12)

Hafkin, J.N. (2002). Are ICTs gender neutral? A gender analysis of six case studies of multi donor ICT projects. In *UN/ISTRW Virtual Seminar, Are ICTs gender neutral*?

Hall, P., Ghimire, G. and Newton, M. (2009). "Why don't people use Nepali language software?" *Information Technology and International Development, Volume 5 (1).*

Harris, E. (2004). Building Scientific Research Capacity in Developing Countries. *European Molecular Biology Organization Reports 2004*, 5(1),7-11.

Hussain, S., Shams, S. & Sarfraz, H. (2012). *Dareecha ICT Training Program for Public Schools in Rural Punjab*. CLE-PAN Localization: Lahore.

Hussain, S. & Mohan, R. (2008). Localization in Asia Pacific. Librero, F. & Arinto, B.P. (Eds.) In *Digital Review of Asia Pacific 2007/2008*. International Development Research Centre (IDRC), Canada.

Hussain, S. (2004). Developing Local Language Computing. *I4D Online*, 2(6).

Hussain, S., & Gul, S. (2004). Localization in Pakistan. Localisation Focus: *An International Journal for Localization*, 3(4).

Hussain, S., Gul, S., & Waseem, A. (2007). Developing Lexicographic Sorting: An Example for Urdu. *ACM Transactions on Asian Language Information Processing (TALIP)*, 6(3).

Horrigan, J. (2010). Broadband adoption and use in America. *OBI Working Paper Series No. 1*. Federal Communications Commission.

Hollow, D., & Masperi, P. (2009). An evaluation of the use of ICT within primary education in Malawi. Proceedings of the 3rd International IEEE/ACM Conference on Information and Communications Technologies and Development (ICTD), Doha, Qatar.

Hussain, S., & Mohan, R. (2007). Localization in Asia Pacific. in Librero, F. (Eds.), *Digital Review of Asia Pacific*, 2007-08 New Delhi: Sage Publication.

Hunt, J. (2004).Successful strategies for addressing gender equality issues in programs and projects: What works? *Development Bulletin*, 64, 53-57.

Hunt, J. (2004a). Introduction to gender analysis concepts and steps. *Development Bulletin*, (64), 100-106.

ILO Evaluation Unit. (2000). *Guidelines for the Integration of Gender Issues into the Design, Monitoring and Evaluation of ILO Programmes and Projects*. Retrieved from http://www.ilo.org/public/english/region/eurpro/moscow/news/2000/gmanual9.pdf.

International Telecommunication Union (2011). *The World in 2010, ICT Facts and Figures*. Retrieved from http://www.itu.int/ITU-D/ict/material/FactsFigures2010.pdf

Internet World Stat, Usage and Population Statistics. Retrieved January 2012 from http://www.internetworldstats.com/stats3.htm

ITU. (2010). World Telecommunication/ICT Development Report 2010. Monitoring The WSIS Targets a mid-term review. Retrieved from http://www.uis.unesco.org/Communication/Documents/WTDR2010_e.pdf.

ITU. (2011). World Telecommunication/ICT Development Report 2010. Monitoring The WSIS Targets a mid-term review. Retrieved from http://www.uis.unesco.org/Communication/Documents/WTDR2010_e.pdf.

Jurmi,K., & Wangchuk, S. (2009). Bhutan. In Akhtar, S., & Arinto, P. (Eds.), *Digital Review of Asia Pacific*, 2009-2010(pp152-159) New Delhi: Sage Publication.

Keniston, K. (2004). Introduction: The Four Digital Divides. In K. Keniston and D. Kumar (Eds.) *IT Experience in India*. Delhi: Sage Publishers.

Lambert, R. (1996). *Parsing the Concept of Global Competence, "Educational Exchange and Global Competence*. New York, NY: Council on International Educational Exchange.

Law, N., Pelgrum, W. J., & Plomp, T. (2008). Pedagogy and ICT use in schools around the world: Findings from the IEA SITES 2006 study. Hong Kong: Springer, Comparative Education Research Centre.

Lennie.J.(2005). An evaluation capacity-building process for sustainable community IT initiatives. Evaluation, 11(4), 390-414.

Lewis, M. P. (2009). *Ethnologue: Languages of the World*. **Dallas:** SIL International.

Lindquist, A. E. (2001). Discerning Policy Influence: Framework for a Strategic Evaluation of IDRC-Supported Research. IDRC Evaluation Unit, Ottawa. Retrieved from http://web.idrc.ca/uploads/user-S/10359907080discerning_policy.pdf.

Lindquist, Evert A. (2003) *Discerning Policy Influence: Framework for a Strategic Evaluation of IDRC-Supported Research*. International Development Research Centre (IDRC): Canada.

Marjan, M. A. (2010). Afghanistan. In Akhtar, S., & Arinto, P. (Eds.), *Digital Review of Asia Pacific*, 2009-2010(pp. 129-134) New Delhi: Sage Publication.

Marker, P. McNamara, K. and Wallace, L. (2002). The significance of information and communication technologies for reducing poverty. Department for International Development. London. 64 pp.

MoIT (2000). Pakistan IT Policy. Ministry of IT, Government. of Pakistan.

Morna, Colleen Lowe and Khan, Zohra (2000). "Net Gains: African women take stock of information and communication technologies". Association for Progressive Communication/FEMNET

MoIC (2004). *Bhutan Information and Communications Technology Policy and Strategies*. Royal Government of Bhutan. Retrieved from http://www.moic.gov.bt/pdf/bips.pdf.

Murison, S. (2004). Elements of a Gender Mainstreaming Strategy A fourteen-point framework. The Capacity Development Group.

Naccarella, L., Pirkis, J., Kohn,F., Morley,B., Burgess,P., & Blashki,G. *(2007). Building evaluation capacity: Definitional and practical implication from an Australian case study. Evaluation and Program Planning 30 (3), 231-236.*

Neimanis, A. (2001). Gender Mainstreaming In Practice: A Handbook. UNDP RBEC.

Neilson, S., & Lusthaus, C.(2007). IDRC Supported Capacity Building: Developing a Framework for Capturing Capacity Changes. Universalia.

Nikolov, R., & Illieva, S. (2008). A Model for Strengthening the Software Engineering Research Capacity, SEESE'08, Germany.

Neilson, S. (2001). IDRC-supported research and its influence on public policy: knowledge utilization and public policy processes: a literature review. IDRC Evaluation Unit, Ottawa. Retrieved from http://www.idrc.ca/uploads/user-S/10468843590litreview_final1.doc.

Phissamay, P.(2009), Laos. In Akhtar, S., & Arinto, P. (Eds.), *Digital Review of Asia Pacific*, 2009-2010(pp. 241-248) New Delhi: Sage Publication.

Pimienta, D. (2005). Linguistic Diversity in Cyberspace – Models for Development and Measurement, In UNESCO Institute for Statistics (Eds.), *Measuring the Linguistic Diversity on the Internet*, UNESCO.

Planning Commission (1999). *Bhutan 2020: A Vision for Peace, Prosperity and Happiness*. Royal Government of Bhutan. Retrieved from http://unpan1.un.org/intradoc/groups/public/documents/apcity/unpan005249.pdf.

Potter, C., & Brough, R. (2004). Systemic Capacity Building: A hierarchy of Needs. *Health Policy & Planning,* 19(5), 336-345.

Powell, E. T., & Boyd, H. H. (2008). Evaluation capacity building in complex organizations. *Program evaluation in a complex organizational system: Lessons from Cooperative Extension*. New Directions for Evaluation, 120, 55–69.

Prado, D. (2010), "Présence des langues romanes dans la connaissance," presentation made at "Présence, poids et valeur des langues romanes dans la société de la connaissance," OIF, Paris, 30 April 2010.

Primo, Natasha (UNESCO, 2003). "Gender issues in the information society" Retrieved March 2012 from http://portal.unesco.org/ci/en/file_download.php/250561f24133814c18284feedc30bb5egender_issues.pdf

Preskill, H., & Boyle, S. (*2008). A Multidisciplinary Model of Evaluation Capacity Building. America Journal of Evaluation 29(4), 443-459.*

Raihan, A. (2009), Bangladesh. In Shahid Akhtar & Patricia Arinto (Eds.), *Digital Review of Asia Pacific*, 2009-2010(pp. 144-151) New Delhi: Sage Publications India Pvt. Ltd.

Raina, R. (2007), ICT Human Resource Development in Asia and the Pacific: Current Status, Emerging Trends, Policies and Strategies. UN-APCICT. Retrieved from http://www.unapcict.org/ecohub/resources/ict-human-resource-development-in-asia-and-the

Rhee, H., & Riggins, F. J. (2007). *Development of a Multi-Factor Set of Country-Level ICT Human Resource Capacity Indicators. UN-APCICT*. Retrieved from http://www.unapcict.org/ecohub/resources/development-of-a-multi-factor-set-of-country-level.

Ryder, Martin. (2006). The Digital Divide. from *Encyclopedia of Science, Technology, and Ethics*. Macmillan Reference USA 2001-2006, Retrieved December 2011 from http://carbon.ucdenver.edu/~mryder/dig_div_este.html

Rural Poverty Portal. Retrieved from http://www.ruralpovertyportal.org/web/guest/country/home/tags/bangladesh

Rahman, Tariq. (2003). "Language policy, multilingualism and language vitality in Pakistan". Retrieved December 2011 from **http://www.apnaorg.com/book-chapters/tariq/**

Sabatier, P., & Jenkins-Smith, H. (Eds.). (1993). Policy Change and Learning: An Advocacy Coalition *Approach*. Westview Press.

Shalkwyk, J., Thomas, H., & Woroniuk, B. (1996). *Mainstreaming: a strategy for achieving equality between women and men*. SIDA Department for Policy and Legal Services, Stockholm.

Scenic Sri Lanka (2007). Retrieved from http://www.scenicsrilanka.com/people-of-sri-lanka.html

Shams, S., Hussain, S., & Mirza, A. (2010). Gender and Outcome Mapping, In Belawati, T. and Baggaley, J. (Eds.), *Policy and Practice in Asian Distance Education*, New Delhi: Sage Publication.

Sorasak, P., & Konsona, C. (2009), Cambodia. In Akhtar, S. & Arinto, P (Eds.), *Digital Review of Asia Pacific*, 2009-2010 (pp.167-174) New Delhi: Sage Publication.

Sood, A. (2002). *Guide to ICTs for Development*. Bangalore: The Centre for Knowledge Societies.

Science, Technology and Environment Agency (2006). *National Policy on Information and Communications Technology (ICT)*. Prime Minister's Office (PMO), Government of the Lao PDR.

UN–APCICT, (2007). *United Nations Asian and Pacific Training Centre for Information and Communication Technology for Development*. Retrieved from http://www.unapcict.org/member-countries/indonesia

UNDP (2003), "Information and communication technologies (ICTs) for Development in National Human Development Reports"

Odame, H. H. (2005). "Gender and ICTs for Development: Setting the Context." In Gender and ICTs for Development: A Global Sourcebook. Amsterdam: KIT (Royal Tropical Institute), 13-24**.**

UNDP. (2001). World report on human development 2001. United Nations Development Programme. De Boeck University for UNDP, Brussels, Belgium

United Nations Information and Communications Technologies Task Force. (2003). Working Group on National and Regional e-Strategies. Retrieved from http://cap.ic.gc.ca/docs/ICT_e.pdf.

Unger, B. & Robinson, N. (2007). .kh Cambodia. In Librero, F (eds) *Digital Review of Asia Pacific 2007—2008*, 122-130. Sage Publications.

UNESCO. (2007). *ICT and literacy*. Retrieved from http://www.unescobkk.org/?id=1980.

UNESCO. (2008). *Improving the Quality of Mother Tongue-based Literacy and Learning: Case Studies from Asia, Africa and South America*, UNESCO, Bangkok, Thailand.

UNESCO. (2010). *Reaching the Marginalized*. UNESCO, Paris, France, and Oxford University Press, Oxford, UK.

United Nations. (2006). *Millennium Development Goals*. Retrieved from http://www.unmillenniumproject.org/goals/gti.html.

United Nations. (1997). *Report of the Economic and Social Council for 1997* (A/52/3). United Nations, New York.

United Nations. (1995). *Beijing Declaration and Platform for Action*. United Nations, New York.

USAID(2011). *USAID Gender Integration Matrix*. Retrieved From http://transition.usaid.gov/policy/ads/200/201sac.pdf.

View of Cambodia. (2011), Retrieved from http://www.cambodianview.com/khmer-language.html

Vose, B.P. & Cervellini, A. (1981). Problems of scientific research in developing countries. *IAEA Bulletin*, 25(2), 37-40.

Web Technology Survey. Retrieved December 2011, from http://w3techs.com/technologies/overview/content_language/all

Weiss, C. (1977).Research for Policy's Sake: The Enlightenment Function of Social Science Research. *Policy Analysis*, 3 (4), 531-545.

Webber, D. J. (1991). The distribution and use of policy knowledge in the policy process. *Knowledge & Policy*, 4 (4): 6-36.

Wattegama, C. & Soehardjo. J. (2011). ICT Sector Performance Review for Indonesia. Retrieved from SSRN: http://ssrn.com/abstract=2013701.

Weerasinghe, R and Silva,C.(2010), Sri Lanka. In Akhtar, S. & Arinto, P (Eds.), *Digital Review of Asia Pacific*,2009-2010(324-334)New Delhi: Sage Publications India Pvt. Ltd.

Wibberley, C., Dack, L. M. F., & Smith, M. (2002). Research-minded Practice in Substance (mis) Use Services. *Journal of Substance* Use, 7(1), 19-23.

Wignaraja, K. (2009). *Capacity Development: A UNDP Primer.* Retrieved from http://content.undp.org/go/cms-service/download/asset/?asset_id=2222277

Wing,T.K. (2004). Assessing the effectiveness of Capacity-Building Initiatives: Seven Issue for the Field. *Nonprofit and Voluntary Sector Quarterly* 33(1), 153-160

World Bank. (2002). Information and Communication Technologies: A World Bank Group Strategy. Washington: World Bank.

Wolff, H. Ekkehard. (2005). The Language Factor in Discourse on Development and Education in Africa. Paper presented to the Symposium on Language for Development in Africa, held at Moi University, Eldoret, Kenya, June 1-3, 2005. (Unpubl.)

World Bank (2005) "Education Notes: In Their Own Language…Education for All", Retrieved from http://siteresources.worldbank.org/EDUCATION/Resources/EducationNotes/EdNotes_Lang_of_Instruct.pdf.

World Bank () *Checklist for the Planning, Design and Implementation of an ICT Project Incorporating Gender Issues.* Retrieved from http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTGENDER/EXTICTTOOLKIT/0,,contentMDK:20271939~menuPK:562595~pagePK:64168445~piPK:64168309~theSitePK:542820,00.html.

Woldemikael, T. M. (2003). Language, Education, and Public Policy in Eritrea. *African Studies Review, Vol. 46, No. 1,* pp. 117–36.

Wikipedia: Retrieved from. http://en.wikipedia.org/wiki/List_of_countries_by_English-speaking_population

Young, J. (2005). Research, Policy and Practice: Why Developing Countries are Different. *Journal of International Development,* 17, 727–734.

Appendix A

Questionnaire – Organizational Capacity

General Information

Respondent's Name: Gender   Country Partner Institute Name:

Organizational Profile

How many research areas/units have been identified within the organization?
How many LLC projects has your organization initiated in the last 3 years?
2 & below   3 - 5   6 & Above
What is the average **annual** funding amount (in dollars) that has been granted to the organization over the last 3 years?
      2009: _____ 2008: _____ 2007: _____
What is the average duration of the projects being conducted by the organization?
      2009: _____ 2008: _____ 2007: _____
      (You may use the scales; Less than 6months, between 6 – 24 months, above 24 months)
Primarily in which of the following LLC domains is the organization working?
      Standardization
      Basic Localization

Language Processing

Script Processing   Speech processing
How many other national and international organizations are collaborating with your organization in context of the PAN Localization project?

Organization Name	Organization Type &	Public \ Pvt.
e.g. ABX Org	Research in Linguistics	Private

Is the research done through PAN submitted/ presented to any other national/ international organization, formally or informally?

Publications (including thesis, MS research, Books, Book Chapters)	Submission Level (Informal/Formal)	Submitted to: Organization Name	Submitted to: Organization Type
Training (trainings attended, conducted, workshops attended, workshops conducted)			
Software Development			

Project Staff Profile

What is the total strength of the research staff working on LLC?
Group them according to their academic levels: *Undergrad & below   Graduate   PhD or Higher*
What is their number per areas of expertise?
Basic Localization   Language processing   Script Processing Speech Processing
Please specify average research publications per employee over the last 3 years
      ☐Below 5   ☐ 5 to 10   ☐ 11 to 20   ☐ 21 and above
What is the average qualification of the organization's leadership?
      ☐Undergrad ☐Graduate ☐PhD. and Higher
At what levels within the LLC projects is the leadership of the organization involved?
      ☐Project Mgmt ☐Project Design   Policy level
Is the organization's staff enrolled on relevant discussion group/communities/Forum?
      ☐ Yes   ☐ No

If yes, please name a few:

How would you prorate the ability and skill development/enhancement of the organization's researchers during the last 3 years?

1 (Challenged)   2(Fair)   3 (Average)   4(Good) 5(Excellent)

Skill Development	Start of Project, Early 2007	Towards project End, Mid 2009
LLC Project development		
LLC Project design		
Problem identification		
Project implementation		
Ability to do analysis		
Ability to communicate results to make research matter		
Ability to research differently (multi disciplinary research)		
Quantitative analytical skills		
Qualitative analytical skills		

Organizational Resources

How far has the project contributed to build organizational infrastructure?

Total Equipment Contribution _____

Total S/W Contribution _____

Linguistic Resources: _____

Books\Library _____

Support of Recurring Expenses? _____

Research Work

How frequently is the staff presenting research at forums? *(Tick any one of the following)*

      Less than 50% staff, representing once a year

      About 50%staff, representing once a year

      More than 50% staff, representing more than once a year

Are there LLC relevant seminar/ training organized by the organization?

      ☐ NO   ☐ Yes   (If yes, then please provide the following)

	Event Name	Event Type (Local \International)

Does your organization belong to any research network?   ☐ Yes   ☐ No

How does the organization benefit from such network?

      Access to technical support: Increase in quantity and quality of research: Other:

Appendix B
Country-Wise list of Research Publications through PAN Localization Project

Bangladesh

Khan Md. Anwarus Salam, Mumit Khan and Tetsuro Nishino, "Example Based English-Bengali Machine Translation Using WordNet", TriSA 2008, Japan.

Md. Abul Hasnat, Muttakinur Rahman Chowdhury and Mumit Khan, "Integrating Bangla Script Recognition Support in Tesseract OCR", Proc. of Conference on Language and Technology 2009 (CLT09), Lahore, Pakistan, January 22-24, 2009.

Md. Abul Hasnat and Mumit Khan, "Rule Based Segmentation of Lower Modifiers in Complex Bangla Scripts", Proc. of Conference on Language and Technology 2009 (CLT09), Lahore, Pakistan, January 22-24, 2009.

Md. Abul Hasnat and Mumit Khan, "Elimination of Splitting Errors in Printed Bangla Scripts", Proc. of Conference on Language and Technology 2009 (CLT09), Lahore, Pakistan, January 22-24, 2009.

Firoj Alam, S. M. Murtoza Habib and Mumit Khan. Text Normalization System for Bangla. Conference on Language and Technology 2009 (CLT09), NUCES, Lahore, Pakistan, January 22-24, 2009. [poster]

Altaf Mahmud, Kazi Zubair Ahmed and Mumit Khan, "Detecting Flames and Insults in Text", Proc. of 6th International Conference on Natural Language Processing (ICON-2008), CDAC Pune, India, December 20 - 22, 2008.

Farhana Faruqe and Mumit Khan. "BWN - A Software Platform for Developing Bengali WordNet", International Joint Conferences on Computer, Information, and Systems Sciences, and Engineering (CISSE 08), December 5 - 13, 2008.

Firoj Alam , S. M. Murtoza Habib and Mumit Khan. "Acoustic Analysis of Bangla Consonants", Spoken Language Technologies for Under-resourced language (SLTU'08), Hanoi, Vietnam, May 5 - 7, 2008.

Indonesia

Adriani, Mirna, Ruli Manurung, and Femphy Pisceldo. Statistical Based Part Of Speech Tagger for Bahasa Indonesia. Workshop, Co-located Event ACL-IJCNLP 2009. Singapore, August 1, 2009. Third MALINDO International MALINDO.

Budiono, Hammam Riza, Chairil Hakim. Resource Report: Building Parallel Text Corpora for Multi-Domain Translation System. 7th Workshop on Asian Language Resource, ACL-IJCNLP 2009, Singapore, August 2009.

Mongolia

A.Altangerel, Journal Scientific Transactions. A Design and Implementation Mongolian Speech Recognition System. MUST № 5/102, 2008.

J. Purev and Odbayar. Corpus Building for Mongolian Language. The 6th Workshop on Asian Language Resources, 2008.

Purev Jaimai and Odbayar Chimeddorj. Corpus Building for Mongolian Language. Proceedings of the 6th International Workshop on Asian Language Resources (ALR)- Jan 11-12, 2008, Hyderabad, India.

Purev Jaimai, Tsolmon Zundui, Altangerel Chagnaa, and Cheol-Young Ock. PC-KIMMO-based Description of Mongolian Morphology. International Journal of Information Processing Systems, Vol. 1 (1), pp. 41-48. (2007).

Purev Jaimai and Odbayar Chimeddorj. Part of Speech Tagging for Mongolian Corpus. 4th International Joint Conference on Natural Language Processing. (IJCNLP). The 7th Workshop on Asian Language Resources. August 2-7, 2009, Singapore.

Purev Jaimai and Odbayar Chimeddorj. Resources for Mongolian Language. Proceedings of the 3rd International Universal Communication Symposium. December 3-4, 2009, Tokyo, Japan.

Nepal

Bal Krishna Bal.Towards Building Advanced Natural Language Applications – An Overview of the Existing Primary Resources and Applications in Nepali. Proceedings of the 7'th Workshop on Asian Language Resources, Association for Computational Linguistics, Suntec, Singapore, August, 2009, pp.165-170.

Pakistan

Hussain, S., Gul, S., Waseem,A. Developing lexicographic sorting: An Example for Urdu. In ACM Transactions on Asian Language Information Processing (TALIP), Volume 6 Issue 3, 2007.

Hussain, S. Resources for Urdu Language Processing. In the Proceedings of the 6th Workshop on Asian Language Resources, IJCNLP'08, IIIT Hyderabad, India, 2008.

Hussain, S., Karamat N., Mansoor, A. Arabic Script Internationalized Domain Names. In the Proceedings of the CIIT Workshop on Research in Computing, CWRC'08, CIIT Lahore, Pakistan, 2008.

Sana Shams. Gendered Outcome Mapping. Outcome Mapping Learning Community Newsletter 2009. No.2.

Sana Shams, Sarmad Hussain and Atif Mirza, Gender and Outcome Mapping. Published in PANdora Distance Education Guidebook (1st edition), 25th October 2008. Research Publications

Sri Lanka

Ruvan Weerasinghe, Asanka Wasala, Viraj Welgama and Kumudu Gamage. Festival-si: A Sinhala Text-to-Speech System. Proceedings of Text, Speech and Dialogue, 10th International Conference, TSD 2007, Pilsen, Czech Republic, September 3-7, 2007.

Ruvan Weerasinghe, Asanka Wasala and Samantha Mathara Arachchi. Facilitating Information Accessibility for the Print Disabled. Diriya 2007 - a Conference on "Mainstreaming Disability into Development". Colombo, Sri Lanka.

Asanka Wasala, Ruvan Weerasinghe. EnSiTip: A Tool to Unlock the English Web. 11th International Conference on Humans and Computers, Nagaoka University of Technology, Nagaoka, Japan, 20-23 November 2008.

Ruvan Weerasinghe, Asanka Wasala, Dulip Herath and Viraj Welgama. NLP Applications of Sinhala: TTS & OCR. 3rd International Joint Conference on Natural Language Processing. (IJCNLP). Exhibitions & Demonstration Session. January 7-12, 2008, Hyderabad, India.

Harsha Wijayawardana, Asanka Wasala, Ruvan Weerasinghe and Chamila Liyanage. Implementation of Internet Domain Names in Sinhala. International Symposium on Country Domain Governance. Nov, 20-22, 2008, Nagaoka, Japan.

Silva, A. M. and Weerasinghe, A. R. Example Based Machine Translation for English-Sinhala Translations. In Proceedings of the 09th International IT Conference (IITC 2008), Colombo, Sri Lanka, 27-28 October 2008.

Asanka Wasala, Ruvan Weerasinghe, Randil Pushpananda, Chamila Liyanage and Eranga Jayalatharachchi . An Open-Source Data Driven Spell Checker for Sinhala. e-Asia 2009. Colombo, Sri Lanka, 2-4 December 2009.

Weerasinghe, A. R., Liyanapathirana, J. U., Asanka Wasala, Dulip Herath, Viraj Welgama . OpenTM: A Translation Memory System for Complex Script Languages. International Conference on Machine Translation Twenty-Five Years On, Bedfordshire, UK, 21-22 November 2009.

Ruvan Weerasinghe, Dulip Herath and Viraj Welgama. Corpus-based Sinhala Lexicon. 4th International Joint Conference on Natural Language Processing. (IJCNLP). The 7th Workshop on Asian Language Resources. August 2-7, 2009, Singapore.

Ruvan Weerasinghe, Asanka Wasala and Kumudu Gamage. A Rule Based Syllabification Algorithm for Sinhala. 2nd International Joint Conference on Natural Language Processing (IJCNLP-05), Jeju Island, Korea, 2005.

Asanka Wasala, Ruvan Weerasinghe and Kumudu Gamage . Sinhala Grapheme to Phoneme Conversion and Rules for Schwa Epenthesis. In Proceedings of the COLING/ACL on Main Conference Poster Sessions (Sydney, Australia, July 17 - 18, 2006). Annual Meeting of the ACL. Association for Computational Linguistics, Morristown, NJ, 890-897.

Appendix C
Questionnaire – End Users

Section A-General Information

Name: Gender: Age: Mother tongue:
Language (spoken at home)   (spoken at work)
Please specify your *education level*?
    Primary (1-5 grades)   Secondary (8-10grade)   Higher Secondary (11-12grade)   University Level
    (13 grades onwards)   Not Educated   Other:
Please specify your *social status*?
    Business/self employed *(if yes, of what type?)* Employed: ♣Government sector   ♣Private
    sector   ♣Social sector
    Student
    Housewife
    Any other (Please specify)

How would you rate your reading	In English	Local Lang.
Excellent		
Very Good		
Good		
Fair		
Poor		
WRITING	In English	Local Lang.
Excellent		
Very Good		
Good		
Fair		
Poor		

Section B (This section is only for those respondents who had prior knowledge of using computers)

Computer Background
For the following, please provide the relevant information according to the skill level mentioned below:
I don't know what this is.
I know what this is but have never used it.
I have used this occasionally but need more training
I use this regularly and feel confident
I would be comfortable showing others how to use this.

Are you trained on the following?	Name of Software Learnt	Language of the Software Learnt (English or Local Language)	Skill level: 1 – 5
General skills: Using Mouse, Keyboard, Printer,			
General file management skills: Opening, saving, copying files , Creating,			
Word processing			
Spreadsheet			
Presentation			

Graphics			
Email			
Internet (Searching)			
Chatting			
Website Development			
Content Management			
Blogging			

Section C: General Computer Usage

How long have you been using the Internet?
    ♣ Less than 6 months   ♣ Almost 2 years   ♣ More than 2 years
Is the English language interface of the computer ever hampered your computer usage?
    ♣Yes   ♣No
How often do you use computer?
    Daily   At least once a week   At least once month   Less than once a month
Where do you usually use a computer? *(Mark as many as apply)*
    At work   At school   At home   At the telecenter   Other
Do you use computer with English interface or computer in a local language interface?
What purpose do you use a computer for?

Mark as many as apply	Give examples of the Purpose, e.g. songs, movies, doing
Entertainment	
Studies	
Work	
Household	

Section D: Using Office Applications

(This section is <u>only</u> for those respondents who have received training on office applications)
    Which office applications are you frequently using?

Name the Software	Language of the interface English / Local Language	Frequency: (Daily, At least once a week, At least once month, Less than once a month, Never)

Which of the following tasks are you able to do in your word processor? *(Mark as many as apply)*
1. Open and close a document
2. Create and save a document.
3. Use formatting options, right left alignment, and L-R directions for languages.
4. Insert logos, pictures in a document
5. Do spelling check
6. Correct errors with backspace and delete keys.
7. Print document.
8. Draw a table

In which language do you mostly develop the office documents?
    English   Local language (in Local script)   English and local language combined   Local language (in Roman script)
For what purpose do you most often use word processor?

Section E: Using Email for Communication

190

(This section is <u>only</u> for those respondents who have received training on localized email)

Do you have your own email address?   ♣Yes   ♣No

    If yes, what is your email address?

Do you prefer to receive your emails in English or in any other language?

    English   Other: (please specify)

To whom do you generally send an e-mail?

    Friends   Teacher/Trainers   Family   Other

How often do you write an email?

    Daily   At least once a week   At least once month   Less than once a month   Never

In which language do you mostly write your emails?

    English   Local Language (in local script)   English and local language combined   Local language (in Roman script)

Indicate your skill level to do the following?

I don't know what this is.

I know what this is but have never used it.

I have used this occasionally but need more training

I use this regularly and feel confident

I would be comfortable showing others how to use this.

TASK	SKILL LEVEL (1 - 5)
Create and send messages	
Attach files to outgoing messages	
Open files attached to incoming e-mail messages	
Attach a signature to messages	
Other	

Section F: Working with Websites

(This section is <u>only</u> for those respondents who have received training on using localized Web Browser)

Which web browser do you usually use to access websites? (Mark as many as apply)

    Internet Explorer   Mozilla   Firefox   Other (please specify)

Which of the following are you able to do in your browser? *(Mark as many as apply)*

    Type a web site address
    Set the home page
    Open additional browser windows
    Bookmark sites
    Select copy and paste text into other applications
    Print a web page
    Upload a web page
    Develop a relatively enhanced web page having links, buttons, formatted text & images
    Save pages to disk

What sorts of Web sites have you visited in the last month? *(Mark as many as apply)*

    Educational
    Games
    Entertainment
    News
    Religion
    Sports
    Other (please specify)

Are there any specific websites that you visit regularly?

    ♣Yes   ♣No If yes, please list a few:

Have you developed any website/websites?

    ♣Yes   ♣No If yes, please list a few:

Evaluation Findings of PAN Localization Project

Website Address (www…)	Description of the website (Newspaper, My school, etc.)	Language (English, Local language in local script, Local language in Roman script)

Please indicate which of the following search engines/services do you usually use?
        Alta Vista   Yahoo   Google   Other (please specify)
What type of information do you mostly like to search through the Internet?
        Work/business   Learning/teaching   Personal needs   Other
In which language do your mostly type your search query?
        English   Local Language   Both
Which of the following are you able to do in your browser? *(Mark as many as apply)*
        Access a search engine on the internet
        Enter appropriate key words into the search engine to locate the desired information
        Refine a search depending on outcomes of the original search
        Use advanced search features, provided in search engine

Content Management Systems

(This section is <u>only</u> for those respondents who have received training on using localized content management systems)
Name any Content Management System that you have used?
Describe few actions/ tasks that you are able to perform using the CMS?
What are the benefits of having a localized CMS?

Chatting

Which chat software do you mostly use?
How often do you chat?
        Daily   At least once a week   At least once month   Less than once a month   Never
Have you ever used localized instant messaging software?
        ♣Yes   ♣No
Which language interface would you now continue to use?
        ♣English   ♣Local
Where did you use localized Instant messaging software? (Mark as many as apply)
        At work   At school   At home   At telecenter / community center   Other
Which of the following tasks are you able to perform in chatting software? (Mark as many as apply)
        Sign in and out to chat room   Visit different chat rooms   Add contacts to the account.
        Chat with the contacts   Change your status during chatting Send and Receive files
        Block contacts   Chat while being invisible
With whom do you generally chat?
        Friends   Teacher/Trainers   Your community   Other
In which language are you more comfortable to chat?

Section G

How would you rate your current skill level in using the software on which you have received training now?
        I don't know what this is.   I know what this is but have never used it.
        I have used this occasionally but need more training   I use this regularly and feel confident
        I would be comfortable showing others how to use this.

Check those software	Name of	Language of the software learnt	Skill level: 1 – 5
General skills:	X		
General file management	X		
Word processing			

Spreadsheet			
Presentation			
Graphics			
Email			
Internet (Searching)			
Chatting			
Website Development			
Content Management			
Blogging			

Which software do you like the most?

Basic Computer Skills   Operating System   E-mail   Instant Messaging (Chat)
Internet Search   Word Processor   Graphics   Web Page Development
Content Management System

How confident are you that you could teach localized software to others?

Totally confident   Very Confident   Fairly Confident   Only a little confident
Not at all confident

How would you describe your overall experience with using localized software?

Excellent   Good   Average   Below Average   Poor

How do you think having local language software can benefit those who did not have computer training before?

Would you be interested in receiving further training on other localized software?

Yes
No
If yes, what type of computer training would you be interested in?

Appendix D
Questionnaire – Policy Development

General Information

Respondent's Name: Gender: Country Partner Institute Name:
Are you a member of any public policy making body?
♣Yes   ♣No

If yes, please provide the following information for EVERY committee (for membership in multiple committees)
Name of the Committee/Authority:
Membership type:
Duration of Membership:
Sector representing:
Name any salient policy issues/standardizations discussions completed
Name any salient policy issues/standardizations discussions in progress
As a member of policy making body describe challenges faced in developing and executing policy/standards development for Local Language Computing
Background Information: "National Policy Documents"

What is your **Language policy**? (Please provide English translation of relevant sections of Language Policy/other relevant public documents referring to the need for LLC)
If available please also provide the following information:
Which authority is involved in the policy development process?
Who are the key players (with designations) involved in the process?
What is your **IT policy**? (Please provide an English translation of the relevant sections of IT Policy/ICT policy/ relevant public documents referring to the need for LLC).
If available please also provide the following information:
Which authority is involved in the policy development process?
Who are the key players (with designations) involved in the process?
Do the above mentioned policy documents directly/ indirectly stipulate initiation of R&D in the respective domain?
♣Yes   ♣No

At which of the following international standards development bodies is your country represented *(you may tick multiple options)*?
♣ISO   ♣Unicode   ♣ICANN   ♣GAC   ♣Other:

Background Information: Policy Development Process

Please list which committee/authority is involved in the following:
Standards development for Language :
Standards development for IT/ICT:
Authorization/Approval of LLC projects:
Has there been any change in the policy development process lately in terms of any of the following:
Membership at policy development forums
Inclusion/Exclusion of a certain Org.
Inclusion/Exclusion of a certain Criteria
Representation
Other: Please provide brief description *(if required)*
Has the project explicitly contributed through advocacy/or any other means to influence any of the changes mentioned in pt. 2   ♣Yes   ♣No
If yes, then please mention which one:
Has the project otherwise/implicitly contributed directly or indirectly in policy decisions/ decision making process?
♣Yes   ♣No
If yes, then please mention how:
Have any of the project outputs directly translated into policy decisions/ laws/ regulation?

☐ Yes ☐ No
If yes, please describe which one and how?

Output Name	Role in influencing Policy

Are any of the following indicators collected by the local statistical bureaus?

Language
Technology
Technology & Language

Are there any inter organizational collaborations to develop the following?

Local Language Resources   Language Technology (software)?

Is Local Language computing curriculum nationally recognized as a coursework to be included in Universities?

♣Yes   ♣No

If Yes, from the options below, which disciplines are currently being taught?

♣Linguistics   ♣Computational Linguistics   ♣CS (NLP)

At what level are the courses taught:

♣Under Grad.   ♣Graduate   ♣PhD

LLC Projects

Are there any LLC projects initiated nationally or internationally that may be seen as a contribution of the project's advocacy/awareness raising/capacity building at public organizations? ♣Yes ♣No

If yes, then please provide the following information:

Title	Granted by	Granted to	Duration	Approx. Funding	Output

Is there any public sector interest in developing more local language computing projects?

☐ Yes ☐ No

Are there any action plans to give out local language computing projects?

☐ Yes ☐ No

What is the primary project's licensing schemes being followed in the public sector?

Licensing	Software	Content

PAN Localization Project's Role in Policy Development

How has the project contributed to raise awareness of decision-makers and other relevant stakeholders about the **potential and impact of LLC**? Tick all mediums used to perform the above:

Seminars and workshops
LLC awareness-raising sessions for policy makers
Developing the capacities of senior and middle level government official
Consultation occurred in drafting the Policy
Assisting government official by providing information
Training requirements associated with the development, implementation or monitoring of the policy
To share best practices
Other

How has the project contributed to raise awareness of decision-makers and other relevant stakeholders **about IPR issues**? Tick a few mediums used to perform the above:

Seminars and workshops
LLC awareness-raising sessions for policy makers
Developing the capacities of senior and middle level government official
Consultation occurred in drafting the Policy
Assisting government official by providing information
Training requirements associated with the development, implementation or monitoring of the policy
To share best practices

Has there been any collaborative work initiated through the project to provide recommendations for public policy development?   ☐ Yes   ☐ No

How would you rate the capacity of your organization to deliver relevant policy recommendation?

Very Limited   Limited   Satisfactory   High   Very High

Appendix E
Gender Mainstreaming Tool of PAN Localization in Localization Projects

		RATING			
		Sub Question (X.Y.Z…)	Main Question	Phase-Wise	Overall
1	Project Design and Redesigning				
1.1	Gender equality objectives have been developed				
1.2	Both men and women have participated in setting the gender equality objectives				
1.3	Gender and ICT issues have been identified				
1.3.1	Gender and ICT issues regarding end-user training have been identified				
1.3.2	Gender and ICT issues regarding localized technology development have been identified				
1.3.3	Gender and ICT issues regarding localized content development have been identified				
1.4	Target population of the project is gender- balanced				
1.4.1	Specific interventions are designed to include participation of both men and women				
1.4.2	Selected target population includes equal participation of both gender				
2	Project Preparation				
2.1	Institutions delivering services under the project are gender-aware				
2.1.1	Gender expert is included in the project team				
2.1.2	Project design team or the implementation staff has received a gender training				
2.1.2.1	Project design team or the implementation staff has received gender training before project initiation				
2.1.2.2	Project team has organized gender awareness seminars during the project				
2.1.2.3	Project design team has attended gender training organized by other organization				
3	Project Implementation				
3.1	Project strategies have helped raise women participation in the project				
3.2	Specific actions that target women have been taken when it is likely that women would be underrepresented in project activities				
3.2.1	Strategies have been employed to address gender and ICT regarding end-user training on localized				
3.2.2	Actions have been taken to address gender and ICT issues regarding development of localized software				
3.3	Men and women have equitable access to project ICT resources				
3.3.1	Men and women have equitable access to the localized technology developed				

3.3.2	Men and women have equitable access to the developed ICT training material				
3.3.3	Men and women have equitable access to the onsite ICT training opportunities				
3.3.4	Men and women have equitable access to online local language content				
3.4	Partnerships are built to enhance outreach and improve access to resources by women				
3.5	Consultations regarding gender issues are held with key stakeholders				
3.5.1	Consultations on gender issues are held regarding the developed localized technology				
3.5.2	Consultations on gender issues are held regarding the end-user training on localized technology				
3.5.3	Consultations on gender issues are held regarding the localized content development				
3.6	Gender is presented in project reports				
3.6.1	Gender perspective is presented in the project contract\proposal document				
3.6.2	Gender perspective is presented in the project contract\proposal document, country progress				
4	Project Monitoring				
4.1	Project has focused on using gender sensitive framework for monitoring and evaluation of the				
4.2	Separate performance indicators are developed for measuring women and men's progress in the project				
4.3	Project monitoring data is sex-disaggregated				
5	Project Evaluation				
5.1	Gender analysis is included in the TORs of the evaluation team?				
5.2	Project evaluation team is gender balanced				
5.3	Sex disaggregated data, including sex-disaggregated performance indicators are collected and analyzed				
6	Project Management				
6.1	Project management team is gender balanced				
6.1.1	Project team involved in execution of the end-user training on localized software is gender balanced				
6.1.2	Project team involved in development of localized technology is gender balanced				
6.1.3	Project team involved in development of localized content is gender balanced				
6.2	Both male and female staffs are employed at all levels within the team hierarchy?				
6.2.1	Both male and female staff are employed at all levels within the team executing end-user training of				
6.2.2	Both male and female staff are employed at all levels within the team developing localized software				
6.3	Professional development opportunities are equally provided to both men and women				
6.4	Budget is allocated for gender implementation project activities.				

Appendix F

Questionnaire– Gender Integration

General Information

Respondent's Name: Gender: Designation of the Respondent:
Country Partner Institute Name:
Section A: End-User Training

This section is related to those organizations who held\organized any end-users training. Please skip this section if your organization did not held\organize any end-users training.

1: Please give the following information about your team members in the project, specifically about those who were involved in end-user training (including the project design, implementation and evaluation):

Name	Gender:	Qualification	Designation\Role in the project

Was there a gender expert in the project?   1. Yes   2. No

    2.1: If "yes", then please specify any one from the following options:
        Within the project team (hired full time)   3. Both 1&2
    External gender expert/consultant
    2.1(a): Please provide the following information against the above selected option:

Within the project team (hired	Qualification:
External gender expert/consultant	Organization :

3: In which of the following phases of the project, did the team liaise with a gender expert for guidance in gender programming? (*Mark as many as apply*)
☐ During the project design phase   ☐ During the monitoring and evaluation phase
☐ During the implementation phase   ☐ During the reporting/ documentation process

4: Please specify any salient actions recommended by the gender expert that had been incorporated in the project's design, planning, implementation, evaluation and reporting phases:

Recommendations\ Actions	Project phase

5: How many project team members received training on gender issues before this project?

6: Did your organization hold any gender awareness seminars/workshops for the project team, specifically for this project (within the project life cycle)?
    Once   Twice   More than 2 Times   Not Applicable

7: Did the project team members attend any gender awareness seminar\workshop organized by any other organizations/institution?
        1. Yes   2. No
 7.1: If "yes", please provide the following information, otherwise skip this question

Designation of the trainees	Agenda\Topic	Organized by

8: Please provide the following information about the end-users of the project:

End-User Groups	Number of women	Number of men

9: Why did the project select the above mentioned end-user groups, elaborate in detail?

10: Please also describe the reasons for the particular number of women and men in each above mentioned end-user groups:

11: Did your project identify the gender and ICT issues regarding the training program?

1. Yes   2. No

11.1: If yes, how were these gender and ICT issues identified? (*Mark as many as apply*)
- ☐ Based on team opinion being the part of that culture
- ☐ Did a situational analysis before the implementation of the project
- ☐ Identified during the implementation of the project

12: What were the specific gender and ICT issues in all the end-users communities? (*Mark as many as apply*)
- ☐ Lack of control over the financial resources
- ☐ Women's lack of participation in decision making
- ☐ Patriarchal social system
- ☐ Gender imbalanced staff\trainers
- ☐ In appropriate training schedule and timings
- ☐ Gender and cultural barriers that hamper women to engage themselves fully in the technological world
- ☐ High illiteracy level of women in the developing countries
- ☐ Women's limitations in mobility to reach the training and educational canters due to the lack of appropriate transport facilities for them
- ☐ Problems of availability and affordability of software and user support
- ☐ Women have insignificant power over decision making as they are confined by the traditions and constrained by the norms and behaviors in their community
- Any other, please specify

13: What strategies your project employed to address above mentioned gender and ICT issues?

14: Please enlist the strategies which were changed during the life cycle of project, if any?

15: To what extent the employed strategies by the project raised the participation of women in the project?

To very little extent	4. To great extent
To little extent	5. To very great extent
To some extent	

15.1: If "to very little extent" or "to little extent", elaborate what were the constraints?

16: Did any end-users quit from the training program?

1. Yes   2. No

16.1: If "yes", please provide the following information:

Gender wise drop out detail (write the total number)	Reason (why did they quit from the program?)
Women: ….	
Men: ….	

17: Was the training material provided to the trainees?

1. Yes   2. No

17.1: If "yes", please provide the following information:

End-users (groups)	Type of training material:	Why did the project select such type, please briefly elaborate: (e.g. provide online material to facilitate women because they do not have access to internet at home)	Language of the training material: 1. English 2. Local language 3. Both 1 & 2	Why did the project select such form of language, briefly elaborate?(e.g. provided in local language to facilitate farmers due to the lack of education)

18: Did the project make any changes in the training material to facilitate the end-users after the pilot study?

1. Yes   2. No

18.1: If "yes", please elaborate those changes in detail: (e.g. gender-related example were included after pilot study)

9: If you want to share any experience regarding the training material, please write here:

Section B: Technology Development

Evaluation Findings of PAN Localization Project

This section is related to those organizations that developed some software. Please skip this section if your organization did not develop any software application under the PAN Localization project.

     1: Please give the following information about your team members in the project, specifically about those who were involved in software development.

Name	Gender	Qualification	Designation\role in the project

2: Please enlist name of the software which were developed according to following thematic areas:

Thematic areas	Name of the software	Language of the software 1. English 2. Local language 3. Both 1 & 2.
General computer usage		
Content/Information Access		
Content/Information		
Communication		

3: Did the project consider the gender concerns/requirements in the choice of software?
       1. Yes   2. No

4: Did the project take women and men's concerns before the development of software?
       1. Yes   2. No

    4.1: If "no", please briefly explain why?

    If "yes" Please answer the following two sub-questions

4.1(a): What concerns of women and men were identified by the project? (*Mark as many as apply)*
    ☐ Lack of applications like multilingual tools and data base
    ☐ Lack of interface for non-Latin alphabets
    ☐ Lack of graphic interfaces for women
    ☐ Lack of automatic translation software
    ☐ Unavailability of computers, internet and their convergence with other technologies (e.g. radio, television, and print) especially in rural areas
    ☐ Unavailability of affordable and easy to use technologies
    ☐ Unavailability of relevant local language content to use the technology
    ☐ Dominant language (English) used in technology that hinders most of the women from marking its use
    Any other…………………………

4.1(b): How did your project identify the concerns\requirement of both women and men?
    Through the general mainstream of the need of ICTs
    Through the team opinions on the behalf of being the part of that culture
    Through a field visit\situational analysis
    Any other…………………………………………………

5: Did the project design software on the behalf of one of the following options?
    According to gender neutral concept   According to the requirement of both women and men
    According to opinion of team members   Any other……………………………………

6: What specific steps\actions were taken to incorporate gender concerns in the development of software?

7: Did the project organize some training\exercise to test the developed software usefulness?
        1. Yes   2. No

    7.1 If "yes", please give the detail about the trainees and what were the positive and negative outcomes?

Gender	Positive outcomes	Negative outcomes

8: Please briefly explain the challenges that the project faced to incorporate gender concerns in the development of software?

Section C: Content

200

This section is related to those organizations that work on the generation of the content. Please skip this section if your organization did not develop any software application under the PAN Localization project.

1: Who generated the content?

End-users   2. Project Team   3. Both 1 & 2

1.1: If the "end-users", please provide the following information, otherwise move on to question number 1.2.

Gender	Men	Women
Number of men and women who took part in content generation		
Nature of content\information (Mark as many as apply)	☐Agriculture ☐Community ☐Household ☐Health ☐Entertainment ☐Educational ☐Religion ☐Sports ☐News Others………….	☐Agriculture ☐Community ☐Household ☐Health ☐Entertainment ☐Educational ☐Religion ☐Sports ☐News Others………….
Selected language for content generation	1: English 2: Local language 3: Both 1 & 2	1: English 2: Local language 3: Both 1 & 2

1.2: If the project team (*generated the content*)

Name	Gender Male Female	Qualification	Designation\role in the project

2: Did the project identify different content requirements of both men and women before the content generation?

1. Yes   2. No

2.1: If yes, how did the project identify the content requirements of man and women? (*Mark as many as apply*)

☐ Identified on the behalf of being the part of that culture
☐ Did a situational analysis before content generation
Identified through any other situation…………………………………………...

3: Please enlist relevant nature of the content requirement for both women and men (target group).

Gender	Women	Men
Nature of Content\Information (Mark as many as apply)	☐Agriculture ☐Community ☐Household ☐Health ☐Entertainment ☐Educational ☐Religion ☐Sports ☐News Others………….	☐Agriculture ☐Community ☐Household ☐Health ☐Entertainment ☐Educational ☐Religion ☐Sports ☐News Others………….
Preferred language for content access	1: English 2: Local language 3: Both 1 &	1: English 2: Local language 3: Both 1 & 2

4: What strategy the project adapted to make the content be available for the target groups?

Target Groups	Gender	Form of available content\information	Language 1: English 2:Local language

		Online Printed Both 1 & 2	3: Both 1 & 2
	For Men		
	For Women		

5: What challenges did the project team face to generate the gender related content in local language?

6: Did the community access the content generated by the project team?

        1. Yes   2. No

    If "yes", please give the answer of following questions

6.1: Please briefly explain that how did the community access that content? (e.g., online content accessed at telecenter etc):

6.2: To what extent the produced content was accessed by both men and women?

Men	1. To very little extent	2. To little extent	3. To some extent	4.To great extent	5. To very great extent
Women	1. To very little extent	2. To little extent	3. To some extent	4.To great extent	5. To very great extent

6.3: What type of content was usually accessed by the women? *(Mark as many as apply)*

      ☐Agriculture   ☐Community   ☐Household   ☐Health   ☐Entertainment
      ☐Educational   ☐Religion   ☐Sports   ☐News   If any other (please write here)

6.4: What type of content was usually accessed by men? *(Mark as many as apply)*

      ☐Agriculture   ☐Household   ☐Health   ☐Entertainment   ☐Educational
      ☐Religion   ☐Sports   ☐News   If any other (please write here)

Section D: Monitoring and Evaluation

Please fill this section if your organization held monitoring and evaluation of the project.

1: Did the project follow any research framework or methodology (e.g. LFA, GEM, outcome mapping)

               1. Yes   2. No

1.1: If "yes", please provide the following information:

Name of the Research methodology that your project used (e.g. LFA, GEM, outcome mapping)	Purpose of using this methodology (e.g. for Planning, Monitoring , Evaluation)	Please share any specialty of this methodology

2: Had sex disaggregated data been collected by the project?

             1. Yes   2. No

2.1: If "yes", please answer the following two sub questions:

 2.1(a): How did your project analyze the data?

        Through focus on over all outcomes (e.g. 55% of end users are able to use computer)

        Through focus explicitly on men and women outcomes (e.g. 25% of women and 30% of men are able to use computer)

        Both 1 & 2   4. Any other……………………………….

2.1(b) Please briefly explains the reason for choosing the above mentioned process to analyze the data?

Please share any experience\issue regarding gender integration in the project which is not covered\captured by this questionnaire. Thank you for taking the time to complete this questionnaire