

## Briefing Paper:

# Multilingualism and the Internet

*The Internet Society's Briefing Papers are intended to act as an orientation paper for members of the Internet Society on a given topic. They provide background, outline current issues, challenges and opportunities, and provide suggestions as to ways forward. The Internet Society welcomes your comments on this topic. Please send your comments to [pubpol-briefing@isoc.org](mailto:pubpol-briefing@isoc.org).*

## Overview

While the Internet is no longer monolingual, relatively few of the more than 6000 languages of the world are currently used to a significant degree on the Internet. Increasing Multilingualism on the Internet is a multifaceted problem that will probably take time and require concerted collaboration to solve. But it is an important issue that deserves the attention and efforts of those who want to make the Internet an accessible tool for all.

Multilingualism is a fact of life in most countries where many languages coexist within the same country, region, or even in much smaller communities. In some regions, such as the Pacific Islands and Africa, hundreds of languages are spoken within the same country and, very often, most people speak many of the local languages.

At the International level, linguistic diversity is increasingly considered to be an important cultural heritage for humanity that needs to be protected and nurtured as inscribed in the UNESCO Declaration on Cultural Diversity adopted in November 2001. At the national level, linguistic diversity is increasingly considered to be a human right and is protected by an increasing number of constitutions. For example, the post-apartheid constitution of South Africa gives official national status to its 11 indigenous languages, which clearly contrasts with many older constitutions in the continent which do not recognize indigenous languages.

According to some renowned linguists, multilingualism is important for at least the following reasons:

- Economic efficiency and productivity
- The establishment of a democratic society
- The promotion of education
- The respect of human rights<sup>1</sup>.

Unfortunately, multilingualism is not as entrenched in the on-line world (on the Internet) as it is in the real world. The Internet started as a predominantly monolingual medium where English was practically the only language in use. However, as the Internet grew and started to reach populations speaking other languages, linguistic diversity increased, with studies showing that it grew exponentially during the 1990s (which is also the period when the Internet reached much of the developed world).

Internet technology has continuously evolved to accommodate new languages, with diverse scripts and alphabets. However, the content has not followed – linguistic diversity on the Internet has stagnated since the beginning of this century, despite the fact that the Internet continues to reach new people speaking different languages, especially in the developing world. Technical limitations impeding language diversity on the Internet have largely been resolved, but insufficient progress has been made in terms of building linguistic presence: for example, in sub-Saharan Africa, only 2.75% of the web pages targeting the African population use indigenous African languages: the rest of the content is almost exclusively in English and French, which are spoken by less than 10% of the population, even in the countries where they are the official languages.

Linguistic representation on the Internet is essential to development, particularly given the Internet's increasingly important role as a major (if not the major) source of information, depository of knowledge, and communications and business medium of this millennium. Inadequate representation will also have an undeniable impact on access to the Internet. Many people around the world cannot use the Internet today even if they have physical access to it, merely because content and applications in their language are not adequately available. Therefore, the issues of linguistic diversity on the Internet and access to the Internet are inextricably intertwined for much of the world's population.

---

<sup>1</sup> Alexander N., "Linguistic Diversity in Africa in a global perspective", 2008

## Challenges

The reasons behind the slow progress of the presence of many languages on the Internet are quite diverse, but can be classified into technical, economic, social and political issues.

Whilst the technical challenges may have been significant a decade ago, mainly for languages that do not use the Latin alphabet, they are much less significant today thanks to significant progress made during the last decade. The remaining technical challenges concern standards, tools and also technical capacity. Most standards at the core of the Internet, also called Internet protocols, are increasingly being internationalized thanks to the efforts of the Internet Engineering Task Force (IETF)<sup>2</sup>. Similarly, encoding problems, which were the most important impediment for the use of many scripts just a decade ago, are being solved by the introduction of Unicode<sup>3</sup>.

However, in many countries, standardization at the national level has been slower, especially in the developing world. As a result, some standards, such as keyboard layouts, are still lacking, hampering the use of many languages on the Internet. As for the tools, few software tools are truly multilingual and even fewer have been localized to a large number of languages, making it very difficult for users of most languages to access the Internet in their own languages let alone enter information. Last, but not least, insufficient technical capacity is a major bottleneck for many languages.

The biggest challenge, however, is on the content development side. Unless there is enough content in a given language on the Internet, there is little incentive for speakers of that language to use the Internet. In many developing countries there are small-scale, often volunteer-based content development activities in local languages. In the past, these efforts were often led by the diasporas that enjoyed better connectivity and more awareness. Today, as local connectivity improves, more content in local languages is being developed from within the country. However, these efforts have not been sustainable and have not been able to produce a sufficient quantity of up-to-date content and information. Furthermore, most organizations, whether private or public, prefer to publish their information in “international” languages such as English and French and, unfortunately, rarely contribute to developing local content in local languages.

Insufficient economic resources are also probably another reason why languages of the least developed nations are not well represented on the Internet. There is no economic incentive for software developers to produce software tools for languages spoken in these nations, even where those languages are spoken by tens of millions of people, as there is limited demand for those products and, consequently, little prospect of reasonable returns. Local governments and other actors also lack funding to support activities targeting the development of multilingualism on the Internet.

---

<sup>2</sup> Ex. Previously, only ASCII characters could be used in the contents of E-mails and in domain names. ASCII characters include basic Latin characters but ASCII does not support the remaining alphabets and scripts of the world. Today, MIME and IDN standards enable respectively email content and domain names in many more alphabets and scripts of the world.

<sup>3</sup> Unicode is a computing industry standard allowing computers to consistently represent and manipulate text expressed in most of the world's writing systems. The first version of the standard was published in 1991 and it is now used in most major software products and standards.

Some rooted social problems have also had a considerable impact on the development of multilingualism on the Internet. For example, illiteracy is a major impediment to the development of some languages on the Internet. In many African countries, most literate people are not literate in their own mother tongue which complicates the issue.

Lack of political will is also a major impediment in many countries. Experience shows that when a government supports multilingualism, the country's linguistic presence on the Internet is substantially improved. This has been proven for many European languages and also more recently with Chinese, which has become the second most used language on the Internet.

## The Way forward

Continuing to accept linguistic divides on the Internet will have far reaching consequences. This is why not only governments, but also the international community, civil society, academia and all other concerned actors should work to change this prevailing situation. Fortunately, there are many possible actions that these actors can perform to improve the situation, working together in a multi-stakeholder partnership. Most actions that could be taken are local in nature and are, therefore, the responsibility of the local actors, but the international community also has the responsibility to assist.

The Internet Society has been active in the multilingualism space for a number of years, working locally and at the international level. One of the lessons learned from this experience is that all too often piecemeal approaches are taken, resulting in unsustainable change. The Internet Society, therefore, advocates first and foremost looking in a holistic manner at the fundamental enablers of multilingualism to determine how they can be strengthened to create a sustainable multilingual presence on the Internet.

A number of actions can be taken to address some of the challenges outlined above. These are targeted at addressing some of fundamental enablers of multilingualism and should be seen as non-exhaustive:

- **Content** – Attaining a critical mass of content in a given language is essential to attracting new and existing Internet users. The following steps should be considered: identifying and enabling important cultural, historical, linguistic, academic and educational materials and information to be put on-line; securing the commitment of government to place its own information, forms, and other mechanisms for citizen interaction on-line, in local languages; gaining support from local information and content generating stakeholders, including, importantly, the media, for moving content on-line. As users and providers create content, new users will be encouraged to go on-line and new content will be generated, all of which contributes to making a language's Internet presence sustainable.

- **Tools and standards** – Measures to ensure that use of the language on-line is not impeded by technology (or lack thereof). Tools and standards, for example, can significantly simplify the use of a language on the Internet through localization of software tools. Similarly, keyboard standards should simplify text input in some scripts, which is currently complicated by multiple incompatible keyboard layouts. Governments should ensure that tools necessary for the use of local languages are available. National standardization institutions need to take the leadership in ensuring that all necessary standards are developed.
- **Capacity building** – It is easy to underestimate the degree to which capacity building is essential to enabling a sustainable multilingual presence on the Internet. Capacity-related factors, including basic literacy in the given language(s), technical abilities in moving paper, oral, and other content to the Internet, and the more complex areas of tools, standards and content development are critical. Targeted actions will be required to bridge specific gaps that exist in a given community. For example, IT literacy in a given language may be a major issue that needs to be addressed once tools in that language are developed. Similarly, where illiteracy is prevalent, content in non-written form, such as voice and video may need to be encouraged.

The Internet Society believes that, in addition to strengthening the fundamental enablers of multilingualism as suggested above, nations seeking to create sustainable language presence on the Internet should consider a "multilingual road-map" approach that could comprise the following key steps: 1) an Internet presence and opportunity assessment of the language in question; 2) an audit of the economic, technical, societal, and political challenges (some of which are outlined above); 3) resource, stakeholder, and leadership identification; 4) solutions mapping and implementation; and 5) audit and review. Such a road-map approach encourages holistic and systematic thinking about the opportunities, the challenges, roles and responsibilities, so that a realistic appreciation of the importance and complexity of attaining a sustainable multilingual presence on the Internet can be achieved.

Even though multilingualism on the Internet is improving, the Internet Society does not believe that it is expanding quickly enough and many languages and their speakers are still marginalized. This creates linguistic divides that will have societal and political consequences and will fundamentally inhibit economic development and social well-being. As much of the world's communication and business migrates to the Internet, ensuring that access is as straightforward as possible for the next billions of users will be critical: creating an Internet experience in the users' own language will be essential to the growth, evolution, and continued success of the Internet.

## References

### General

Alexander N., Linguistic Diversity in Africa in a global perspective, *Literacy and linguistic diversity in a global perspective – An intercultural exchange with African countries*, 2008, European Centre for Modern Languages

Klensin, J., Issues in Building a Multilingual Internet,  
[http://www.isoc.org/educpillar/resources/docs/multilingual-internet-issues\\_20080408.pdf](http://www.isoc.org/educpillar/resources/docs/multilingual-internet-issues_20080408.pdf)

Securing a place for a language in cyberspace, Diki-kidiri, M.,  
<http://unesdoc.unesco.org/images/0014/001497/149786E.pdf>

UNESCO, Universal Declaration on Cultural Diversity,  
<http://unesdoc.unesco.org/images/0012/001271/127160m.pdf>

Measuring Linguistic Diversity on the Internet, UNESCO,  
<http://unesdoc.unesco.org/images/0014/001421/142186e.pdf>

Cultural and linguistic diversity in the information society, UNESCO,  
<http://unesdoc.unesco.org/images/0013/001329/132965e.pdf>

### Statistics

Top 10 languages on the Internet, <http://www.internetworldstats.com/stats7.htm>

Global Internet Statistics, (by Language), Global Reach,  
<http://web.archive.org/web/20041011043725/http://www.gltreach.com/globstats/index.php3>

Languages on the Asian and African Domains, Language Observatory Project,  
<http://gii.nagaokaut.ac.jp/cdg08/FILES/p4.pdf>