IGF2021 Universal Access and Meaningful Connectivity (UAMC) Main Session: How Can We Achieve a Multilingual Internet?

Report of the Preparatory Session on November 01, 2021

draft 2 - c.a. - 11-Nov-2021

The IGF 2021 program content is based on two baskets: a main focus area basket with two focus areas, and an emerging and cross cutting issue basket with four issue areas. The goal of the preparatory sessions is to refine the scope of these issues with input from the perspective of different regions and stakeholder groups, establishing common ground on expectations for IGF 2021 outcomes.

Most of the preparatory sessions are supposed to be capacity-building-oriented.

The chair of the MAG, Anriette Esterhuysen, presented an overview of the objectives of the preparatory sessions. Wim Degezelle mentioned that the preparatory phase will help to understand how the issue areas were selected, and what are the IGF 2021 goals. These sessions will help people to provide input to the process.

The UNDESA representative described the focal issues of IGF 2021, and how they are linked to and affected by the Sustainable Development Goals (SDGs), and referred to the impact of the pandemic in the implementation of those goals. UNDESA expects IGF outcomes to help advance the SDGs in all countries.

David Souter, managing director of ICT Development Associates, and member of the UK IGF Organising Committee, presented the keynote address, focusing on what he considers are the eight challenges for Internet governance, which affect all aspects of the Internet. His presentation is broad enough to encompass all IGF 2021 focus areas.

He describes three "critical changes" in the Internet ecosystem impacting on Internet governance:

- the growing importance of the Internet in economy, society and culture;
- commercialization and the concentration of economic power;
- proliferation in number and scope of decision-making bodies.

Souter goes on to list and comment on what in his view are the eight challenges:

1. The pace of change and changing nature of the Internet: scale and technology
2. The Internet as part of digitalization (from fixed telephony to the semantic Web)
3. The concentration of digital power
4. Digital geopolitics (and the environment)
5. Shaping the digital future: opportunities and risks (preserving what we value; promoting what we want; preventing what we fear)
6. The future of regulation (permissionless innovation versus the precautionary principle)
7. Multilateralism and multistakeholderism
8. Participation in decision-making

According to Souter, depending on how we approach the Internet (a tech phenomenon; a set of rules and protocols; essentially an economic and cultural phenomenon), we prioritize governance of the technical aspects and leave out all the other aspects. As the Internet advances, new governance issues emerge. Internet governance as it is today is far from sufficient for taking into account the evolution of applications, ranging from pervasive social networks to autonomous vehicles, IoT and AI (which are, in several cases by choice, fundamentally based on Internet protocols).

Big transnationals of digital services (mostly US-based corporations) became so dominant that they swim in the waters of national legislations of most countries practically unbounded.

Digital cooperation, as defended by UNSG, continues to become harder, making it difficult to establish a universality for Internet governance. Development of the Internet should not be guided by the advance of technology.

Souter recalls the motto of the IGF of ten years ago: how to achieve a people-centered, inclusive and development-oriented information society, which requires broad-based and inclusive dialogue, which is something for which the IGF is especially appropriate, both globally and nationally?

He warns against the permissionaless innovation mantra for Internet-related services versus the precautionary principle which generally applies to other economic sectors.

In closing, he mentions the shortcomings to multilateralism and multistakeholderism in decision-making regarding regulations and policies at national and international level pertaining to Internet governance.

Sonia Jorge, executive director of A4AI and head of digital inclusion program at Web Foundation, develops the concept of meaningful connectivity, mentioning the reviews of the ITU/Unesco Broadband Commission, which stressed the challenge of affordability, and recalling that current definitions of universal access are based on a binary concept of online or offline.

In the conceptual schema she presents, three elements make up meaningful access: affordability, social environment, and meaningful connectivity. Meaningful access should essentially facilitatte meaningful use. In her view there are five definition foundations for meaningful access:

- relevant: to enable policy ambitions;
- measurable: to track progress in connectivity;
- evidence-based: to reflect the needs of Internet users;
- gender-responsive: to measure and close the digital gender gap;
- open Internet access: to exclude artificial restrictions.

As to meaningful connectivity, she presents four dimensions:

- sufficient speed
- a smart device
- sufficient data (no data caps)
- sufficient relevance (regular Internet use)

Of special significance is the dimension of sufficient data -- unlimited connection at home or place of work/study, no data caps, no limitations in speed depending on usage.

While affordability and technical aspects of connectivity are measurable, there are no complete standards for measuring meaningful access as a whole, especially because of the social environment elements which include, inter alia, skills, content, language, access rights and so on.

She emphasized the importance of participants bringing to the IGF 2021 main session inputs, cases, stories that illustrate the social elements in an attempt to document them and start developing a clear definition and a measurement framework.

Carlos Baca, general coordinator of Mexico's Research Center in Technologies and Community Knowledge (CITSAC), made a presentation on community networks and their relevance for capacity building, based on experiences of community networking collected by LocNet/Rhizomatica/APC in Kenya, Nigeria, South Africa, Indonesia and Brazil. He presented six points to show the relevance of capacity building in community networking:

- people learn to install, operate, maintain, and manage their networks;
- critical thinking of technologies allows weaving other ways of connecting and interacting with them;
- networks are sustained from and through the political and organizational structures of the community;
- the local economy is strengthened, not deteriorated by the incorporation of technologies;
- the creation and dissemination of local content is encouraged;
- knowledge and information are socialized, resonating in tne development of other community communication and telecommunication projects.

Technologies, in Baca's view, should take into account local dreams and training needs; technologies should be ways of learning and exchanging knowledge; technological experience and expertise should be shared as much as possible; should take into account the economic and social aspects; communities should learn from each other's experiences; paradigms should be broken and fears from technologies mitigated.

He concludes that "technologies must be adapted to the lifestyle and development goals of the communities, not the other way around."

Adriana Labardini, Special Projects and Board of Directors, Rhizomatica (Mexico), brings to discussion the very relevant topic of spectrum allocation, and presents an important critique of the way in which spectrum is allocated by the ITU and national regulators -- granting rights of use of large chunks of spectrum to operators for 20-30 years while blocking innovative ways to optimize spectrum use, such as sharing, secondary use and expanding unlicensed spectrum bands.

Access to spectrum by providers and the community needs to be democratized. A crucial point is the set of reg policies for spectrum sharing at the community/local level. Wi-fi is a successful example of unlicensed spectrum sharing. Spectrum should be freed at local

level for innovative community applications and alternative models of access. The current processes of regulating TV white spaces should bring another example of successful spectrum sharing.

Edmon Chung, CEO of DotAsia Organisation, AprIGF secretariat head, presents the issues related to language in domain names, email addresses and other identifiers. Multilingual techniques have been developed which in practice work for a limited number of languages and still depends on wider acceptance by operators.

This is a challenge being studied, discussed and implemented with difficulties for decades. Chung sees this as a market failure since the end user is not able to fully take advantage of access. The so-called "semantic Web" is still far from bringing a consistent and generalized solution to this challenge.

Chenai Chair, special advisor on Africa Mradi Innovation, Mozilla Foundation discusses the implementation of voice-operated alternatives to communicate with devices and the network -- a major challenge in the great majority of idioms today.

Chenai Chair complements the points raised by Chung and focus on the ability to speak to a device in the user's language instead of having to use a keyboard or similar in a restricted set of idioms. Chenai is concerned that all over the next decade speech is going to become a primary way people interact with devices from laptops to phones and today's voice enabled devices only respond to a small set of idioms.

She exemplifies mentioning that most of the world's languages, language accents and patterns cannot be used by devices such as Amazon Alexa, Apple Siri, Google Home and many others. In addition, those systems tend to work better for men than they do for women. This is quite exclusionary since it means that one has to switch the way they speak or has to find someone who the system understands.

Actually speech recognition data is siloed so you find that only a handful of companies handle the majority of speech recognition interactions, and they often do so using their own proprietary data. So there is a need to tackle this kind of divide in the ability to interact through voice with digital systems.

Common Voice was launched in 2017, aiming at levelling the playing field, while mitigating certain biases. It is an ambitious Open Source initiative with the purpose of democratizing devices and diversifying voice technology. Common Voice actually seeks to enable people to contribute their native voices to a free and publicly available database, so startups, researchers and

developers can use to train voice enabled apps, products and services.

Today, the initiative actually presents the world's largest multi-lingual public domain dataset with more than 9000 hours of voice data and 65 different languages represented. This is an effort really being driven by the community, because it's only possible when people in a particular community gather around to have their language represented on the platform.

So to sum up my point is that at the end of the day, our key goal is to really take in a language intervention that allows the platforms to access open data sets in people's languages, showing that is is possible to develop voice recognition for the language of underserved communities as a platform.