

FNISHED COPY

WORLD SUMMIT ON INFORMATION SOCIETY FORUM
ICT4ALL: INDIGENOUS LANGUAGES MATTER FOR PEACE, INNOVATION AND
DEVELOPMENT
GENEVA, SWITZERLAND
10 APRIL 2019
1500 CET

Captioning Provided By:
Caption First, Inc.
P.O. Box 3066
Monument, CO 80132
800-825-5234

This text, document, or file is based on live transcription. Communication Access Realtime Translation (CART), captioning, and/or live transcription are provided in order to facilitate communication accessibility and may not be a totally verbatim record of the proceedings. This text, document, or file is not to be distributed or used in any way that may violate copyright law.

>> MODERATOR: My name is Irmgarda Kasinskaite-Buddeberg, I'm the programme specialist for the Knowledge Societies Division, Communication and Information Sector of UNESCO. This panel is talking about Indigenous Languages Matter for Peace, Development and Innovation. And I'm pleased to be with you here. And on my left side we have Emily Taylor from Oxford Information Labs, U.K., on my right side we have Eirik Larsen who is political advisor and representative of Sami Parliament, we have Professor Joseph Mariani, laboratory national French scientific Center for research and young fellow Rebecca Forthgrin working with WIPO in Geneva. Welcome to the panel and before I give the floor to the speakers I would like to give you information why we are inspecting about indigenous languages.

Those who are involved in WSIS process, you know that there is one actual line dedicated to cultural and early beginning when we had discussions for Action Plan preparation and during different phases, indigenous issues were taken into consideration at that time and we have several Paragraphs which refer to promotion of indigenous languages and a reflection indigenous

issues in general. So this is why we see an importance to address once again and we have a very good occasion because 2019 is afternoon international year of indigenous languages and UNESCO is produce and takes full responsibility to be a lead UN agency and were glad as well to see what agencies like ITU, WIPO, UNDESA and many other work together with us.

This panel will be dedicated to the discussions how indigenous languages could be mainstreamed, how the issues related to language technology could be addressed in order to facilitate integration of indigenous language speakers around the world. And this is an important subject because UNESCO has been working for a long time analyzing trends of language disappearance around the world. We have languages in danger, which was made available a good decade ago, and we firmly can state today that 40% of world linguistic diversity of languages spoken around the world are in endangered.

There are different levels of endangerment, those which are instinctive and those which are vulnerable. Many are oral languages which are not present in cyberspace which do not have oral, which are based on oral system and don't have origin system. That's why it's important to address issues in the avenues we have today without speaking too long myself, I would like to ask first of all, to share our representative from Sami Parliament what does it mean for indigenous persons speaking indigenous language, Sami language, what does it mean to be, to participate in digital world?

What are the obstacles you see, and what could be the solutions provided by technical community because we know very clearly what Sami Parliament is very much involved in the discussions and promotion of Sami language around the world and you work actively with as well institutions, universities, and as well industry members, so it would be very interesting for us to hear how we at Sami Parliament on discussions what are the strategies that could be used in engaging with industry members?

>> EIRIK LARSEN: So thank you. (Speaking non-English language). Good afternoon, everyone. I am Everitka. My formal name is Eirik Larsen or paper name, if you would like. First of all, I want to bring greetings from the Sami President Ms. Eili Kastkatelo who was not able to participate this afternoon and I guess I'm the substitute. Entering 2019, we are very happy to wish everyone a happy international language year and we have a lot of expectations for this year.

I present with my traditional Sami name, Lasa Everitka which is composed by my grandfather, my father and my own first name. And I chose to do that because that is one obstacle. When I tried to register to this event, I couldn't use Ms. Name. I had to use my paper name. So when we deal with indigenous issues, we

have to rethink the established truths. Even if major languages uses a first name and a surname, there are many languages which name people in another way.

During the colonization the Sami people were faced to take Norwegian, Swedish, Russian name. In Norway you had to have a Norwegian name to own property, but the wind has turned and more and more want to use our traditional name. So I wanted to start with pointing out one of the obstacles, and it's on a basic level to write your own level and to present you with your own name.

I haven't mentioned our letters, we can't write our letters in every software, and if we can, maybe in the printout our letters look very strange. Shall I proceed with the other questions or?

>> MODERATOR: Thank you very much for sharing at least first insight about very important obstacle what something basic for many of us, use of our own name is not possible when you are indigenous languages user and if your language is not very well supported on Internet and this could be an obstacle for indigenous peoples around the world. Maybe I want to ask you another question. Maybe you could share with us how the Sami Parliament is addressing this issue, because we know what Parliament is engagement.

different electronic services and content production.

>> EIRIK LARSEN: I would like to say that in the state of Norway they have in their legislation, new systems, new software, they should handle Sami letters. It will help the letters, but not how to register your name in our tradition. So there are some improvements. The Sami languages, it's, is it nine Sami languages, I think it's nine Sami languages, all of them are threatened. Eight it is.

The situation for some of them are better than others. We have been working in cooperation with university, the arctic university and the Norwegian Government with creating language tools, computer keyboards, mobile phone key boards, dictionaries with grammar, spelling checkers, word Forum generators, grammar checkers, machine translation. And we have actually achieved it for some of the Sami languages.

Our languages have a more complex grammar than the big European languages. I'll say that it's our languages are more advanced and this excludes our languages from the digital modern solutions for the major languages. The solutions are not able to handle our advanced languages, if I can say like that way.

But we have overcome some of the obstacles. We have, as I said, developed all of the tools and you can use Sami letters, grammar checkers, and so on on the office Microsoft, and Apple

and more.

>> MODERATOR: Thank you very much. So it means there is kind of strategy which you will apply in being engaged with differing stakeholders because what you listed clearly what you have are different from many European languages writing and as well oral system, how those languages are being presented so it means that there are different ways how you could as well promote them being used in cyberspace, and you referred as well a collaboration with universities as well as production of special key words, dictionaries, and as well referring to different legislation.

So these are the few things which could be applicable for strategy development.

>> EIRIK LARSEN: First of all, I forgot to mention that all of our infrastructure is open and possible to use for everyone. And there are similar tools in Russia and Canada, and in two weeks in New York during the UN permanent Forum, we have a side event where we will present our solutions for others. And, of course, the best thing is if the major companies, technology companies have integrated the solutions in their software from the beginning, but I think I'm doubtful that that will happen.

So at least we are asking to get access to their source codes so we can develop what we need ourselves and then integrate it ourselves in the systems. And this year happens to be the international year of indigenous languages, so we are trying to develop a momentum among indigenous peoples all over the world to get the major companies to take more responsibilities and we are as well working on an outcome Document of the international year of indigenous languages, and we hope that state commits themselves to work on how to improve the situation for indigenous languages.

And maybe in some states there will be a possibility to make legislation that can secure that indigenous languages are used also in private sector. But I have to mention, speaking about indigenous peoples, there are a lot of variations. Some people are isolated, some people are struggling to get access to electricity, some people have not a written language. Some people haven't Internet and computers and so on. And some are in a very different situation. We as Sami people do different work. We have the responsibility to push our situation so others can follow.

So you have big ambitions on language technology. I have one example. There are lack of teachers and teaching materials in our languages, and they are very expensive to develop, very expensive to develop. Each book is expensive. And the new era with digitalization, it will be a solution for us. We can have digital teaching materials. It's easier to update, and as well we are using teaching through video. And we are now looking into

use robots. Pupil sitting somewhere, maybe the only pupil with Sami, he can participate in another classroom through a robot and participate in the social environment.

So that's the next step we are hoping to get Norwegian Government.

>> MODERATOR: Thank you very much, Eirik. This is why I think it's a very good transit way to professor Mariani. I would like to ask him a few questions about technological developments because this is what was just said by Mr. Eirik Larsen, there are already being robots being used for Sami language and we can see more and more if Artificial Intelligence will be helping if it is integrated at the beginning into the development processes what language would be supported, different language scripts, even those which do not have written systems would be integrated in the design process.

So we are very keen to learn from your expertise as a person who has been working all of their lifelong in computer science as well as specific learned speech processing. So what are the most recent technological developments, innovations which are as discovered by research community and as well by industry which could support multilingualism and as well help speakers like Sami speakers and other indigenous peoples around to access resources, services around the world using electronic means.

Maybe you could share with us some ideas what are the trends and developments?

>> JOSEPH MARIANI: You are right to say that language technologies after incubating in research laboratories for more than half a century, more than 50 years are now being deployed largely in many different applications and for several languages in many different areas. Hour, the situation is very different across the more than 7500 languages that are spoken all over the world. Less than 150 of those languages have language technology.

So that means two persons of the 7500 languages. Less than ten are relatively well covered and only English we may say is largely covered. So the question is what should we do for the remaining 98% of those languages. So some examples of innovation, innovative applications which are now becoming available, I will take two examples. One is the speech interface for the voice, personal voice assistant, and one example is like the Google Home Assistant.

And here it works for certain languages with taking into account language varieties, so, for example, there are 40 varieties for English, so you can find British English, American English, so on, so forth. There is also speaker recognition, it can recognize six different voices so depending on the person in the family with advancing the Google Assistant. There are 150

different voices and also including recently celebrity voices. You can use it for, shopping, and also for booking and these are the complex system that makes it possible the assistant takes an appointment on your behalf or make a booking for a restaurant on your behalf by voice calling the restaurant or the addresser.

Another example machine translation and I would like to take Google Translate. So the characteristics of this system is that it's free, it's online, it's instantaneous, and it's also perfect with a variable level of quality depending on the language. So for the time being it addresses 103 languages which corresponds to roughly 10,000 different language pairs. There is also speech input output for 41 languages with varieties. So certain varieties, for example, and 23 for Spanish, and presently 500 million users are using Google Translate every day, and Google Translates the equivalent of 1 million books per day, which is more than translators do all over the world within a year.

So the challenge of multilingualism is twofold, on the one end it is preserving the culture and the corresponding associated languages, and this is preservation of languages, but at the same time on the other end, it is allowing for communication across languages and this is communication. So both preservation and communication should be taken into account.

And the effect of linguistic technology is double. Language that miss language technologies will be less and less used. If you have to shift from your language to another language whenever you want to use your car navigation system or your Smart Phone or your special vocal assistant for doing Internet search or emergency access, then you will use your native language, your own language. At the same time language that benefit for cross language technologies such as machine translation will be more and more used.

You can think of what would be the difference in the presentation. Presently 96% of the papers in the science are in English, 96%, and focus for remaining languages, so translation will be much more different if there was quality machine translation available.

So native speakers of under resourced languages are, therefore, in an unbalanced situation creating a digital divide and the corresponding languages are in danger of digital extinction if not complete extinction.

>> MODERATOR: Thank you very much Professor Mariani. I still would like to ask you one more question, which probably can as well be provocative. We see very often solutions which are being proposed to the different users and we promote mono lingualism instead of multilingualism. That there are no tools means awareness, what is exactly why solutions cannot be

immediately based on multilingual options, features, and if there is no way what are the obstacles, what can be done, for instance, to use language technology for utilization of languages around the world by communities themselves? So what needs to be done in terms of designing tools which would be serving much more broader community than just only few ones?

>> JOSEPH MARIANI: Maybe we should mention what are linguistic technologies. We can speak first of written language technologies, and here as you say, we need keyboards and keyboards don't exist, may not still exist for all languages. Our basic technology, just a spelling checker, grammatical checker, we have full system for regional language processing, text understanding and generation, automatic summarization. We may speak of search engines, information retrieval, Q and A but also check books, knowledge access, so this is a broad range of such technologies. We go through sentiment and analysis and cross lingual information retrieval when you want access to information whatever the language in this it has been encoded and, of course, automatic machine assisted translation.

Then we go to spoken language processing, and we should speak of speech recognition, and we have here a very nice example here of speech conscription, automatic speech transcription which makes that that many errors even for my very bad English. But when it goes to a different language and you may have noticed previously when our friends from Sa. I started speaking Sami it says non-English speaking so it works for English, but usually not for other languages.

So recognition and understanding, you have speech generation, speech recognition, emotion detection and same language identification and speech translation and interpretation. And not to forgot a third kind of languages, sign languages, and also in terms of analysis, synthesis and translation. I would like to stress the importance and the interesting features of spoken language because spoken language interfaces are especially interesting when with the fact that it only requires simple, cheap, and largely available cell phones and also that it can be used by the illiterates and by the many languages that don't have any writing systems that you mentioned.

So for developing language technologies, it needs to have two different ingredients. The first one is language resources. Most of the systems presently are based on Artificial Intelligence, namely machine learning, and so you need huge amount of data to train those systems. When I say huge, it means big enough words and thousands of hours of speech that you have quality systems. And the more you have for training the better the quality of systems.

This data is also interesting for research in linguistics.

So those language resources are relatively easy to get for some languages. The 2% I mentioned previously, but they are more difficult to get for the remaining 98% of them, and in some cases it may also not be that appropriate depending on the features of the language.

The second ingredient is language technology evaluation so this is assessing quality of the system, so the state of the art and progress, and the data monitor research investigations and antiperformance with the application needs and also to determine whether language technology is usable for a language or for a language pair in terms of machine translation.

For this we have metrics for speech recognition we use what is called word error rate stressing the fact that speech recognition system with 100% word error rate is still a speech recognition system, but is not usable at all. We have the same also for machine translation we use the blue measure which is both the accuracy and fluency of machine translation, and we can compare with human performances, and in very few cases the machine is better than the human for those tasks.

This has been in international evolution campaigns which compare the performances of different systems with different approaches on the same data, on the same test data with the same protocol, but these exist only for very few languages. So the question is now is it necessary to devote the same amount of effort, in order to address the remaining 98% languages given that we only address for the time being 2% of the effort.

So I believe for this we need to know what are the languages, so we need to have a language map, what are the existing languages, the direct variations, populations speaking those languages, the language similarities what are the language resources that exist like newspapers, radio, TV programmes. Then to take into account the peculiarities of the languages, the complex grammar which was mentioned previously or the language which is specific, morphologically rich languages also and we can try different approaches.

One approach is it to do multilingual approach. If you are able to do speech recognition for 20 languages and you want to address 21, one more language, then maybe you can use what you already did in terms of making models of the names of the previous languages to address these two languages if it's not too far for the previous 2021.

We can do language or dialect as well and we can work on the language families all together. It makes things easier. Another approach which is used for automatic speech recognition is what is called the zero source speech recognition. So it's mimicking the way children learn to speak so they don't have any grammar. They learn naturally by themselves by getting examples from their

parents and so they can learn from scratch a language. So maybe mimicking the same for the machine.

Another approach for machine translation is what is called zero approach. So this is based on experiments which were made on developing neural machine translation for Korean to English and vice versa and Japanese to English and vice versa, and from this it was possible to develop machine translation for Korean to Japanese, even so there was no Korean to Japanese. Based on the neural network, and so using the information which was trained which was gained in the neural network, and so we don't translate anymore words or sentences, but we are really translating meaning in that way.

So I would also like to stress finally that the situation is very different across the different languages as was already expressed. In some cases where the language is already extinct or almost extinct we should look at only language documentation for the language which are indigenous languages in order to conduct revitalization of those languages. So it is necessary to produce basic language resources and technologies. For those that are not endangered or not yet endangered, then develop and deploy technologies and applications or adapt applications to those languages which would improve the quality of machine translation technology for most languages and machine translation which is not too bad for some languages but very bad for most of them.

Finally, we should also develop advanced language technology like spoken dialogue systems or robotic interaction which are still open research problems. This is for all languages.

>> MODERATOR: Joseph, I have one more question for you. If we understand correctly, there are different situations where language is at the moment, so we could be those which would require a strong documentation effort because it's not evident what we would be able to ensure the intergenerational transmission in some cases, and these are very sad cases. So what I understand language documentation process is needed to be in place, but there are those languages which could be supported with language technologies.

One thing that we hear very often speaking with industry members, those solutions are quite costly and we are not, we don't produce immediately or even in in the long term economic benefits. So maybe you could elaborate more what could be those economic benefits? What would stimulate industry to produce more language technologies, services content, resources which could be actually used by more communities around the world?

>> JOSEPH MARIANI: I think you are right, it seems that there is a growing interest of companies for languages. Of course, we see many U.S. companies involved in this area,

increasing amount of companies. Also Asian companies are now coming to this business, and much less I must say in Europe surprisingly. So if we look at this story of computer science, computer engineering, it started where companies were just interested by English. Some companies just used English, American English keyboard at the very beginning for those were there at the time as myself, and then they went to major international languages and then gradually to I would say language of economical interest.

There was a time where Microsoft, for example, would make a list of languages based on the national GDP, and then they would go and say I will address languages up to this rank because those are of economic interest. And then more recently, it seems that it went and this is thanks to the social networks and also to the domestic use as I mentioned with Google home, for example, to more languages which where a large population.

There are many speakers, many users, then it's interesting, of course, for those companies which are working in the social network business to have more languages. So now what is the challenge? The challenge is to address more languages so languages, especially languages with low GDP and small population. That's the challenge given that worldwide, 95% of the 7500 languages are spoken by only 6% of the population. So small number of people speaking each of those languages.

And 90 of those languages are disappeared in the next century. Interestingly, there was some big debate very recently in France because the people in Brittany, they were very upset because there was Google Translate for one and not the other, and they were wondering why Corsican and not Britain, so it appears people are eager to have technology in their language as soon as there is technology for people in the same situation as they are.

So I would say that it's a political issue because language technologies, language interfaces are not yet access to function, of course, but to commerce, to knowledge more and more, and soon access to connected objects as well with the coming of the Internet of Things. So this will increase still the necessity of having language interface. It is also ethical issue because we have to get access to personal data to train the systems so it's an issue of privacy. And it's a technical issue because the more you get, the more systems you deploy and the more we get the better will be the systems of a company, it's very important to deploy those systems as largely as possible.

So now the challenge is how to address indigenous minorities original languages both as dialects, migrants, status languages, foreign and regional accents, indigent languages and even extinct languages. As I said, it's both political, cultural, economical, technical challenge. So the questions are should all languages

be equally served? What about languages that are spoken by low GDP countries, what about languages spoken by small populations. What about languages that do not have the chance to be spoken by the U.S. Department of defense as so called peace keeping languages dispatched to languages that are victim of war, tsunami, earthquake, such Creole. What would be the future? What is the cost? Who is ready to pay for it? How can we reduce the cost by using new research directions? So the goal would be to open the language barrier so everybody should have access to language technologies in his or her language, native language including indigenous languages.

So for this, we should, of course, work possible solutions and the purpose would be to pay special attention, of course, to minority languages including indigenous languages and to develop multilingual technologies, so not focusing on single language but more largely on several languages family of languages.

In the European Union, for example, I think it would be important that the European Commission joined the effort with Member States, with also regions and industry to develop technologies for at least the 24 languages, the official languages of the European Union but broader addressing Corsican, Basque and so on and internationally speaking we can extend to large number of companies. Promote language in universal systems as you proposed to the companies talking in the peculiarities of the languages and here I believe UNESCO should play and would play a very important role in coordination.

So we propose in this framework to organize a world Conference called LT for all within the international year of indigenous languages 2019, how to develop language technologies for all people, all indigenous, native minority, regional, territorial, and non-territorial, stateless, endangered languages and dialects and this is how we commit to speak to those items.

>> MODERATOR: Thank you very much, Professor for this intervention. Maybe to add something which was said, I remember a year ago when we were exporting our data set from our older version of languages in danger we made a statistical analysis and we came to very striking results which we actually understood what majority of languages which inscribed very endangered civilian and those extinct are spoken by a small number of speakers and we saw that really less than 10,000 speakers where the intergenerational transmission is core to pass from one again race to another language.

We identified some which were vulnerable, which means most vulnerable which are very close to, and we worked with Oxford Information Labs from U.K., what most vulnerable languages which sometimes we don't see them very closely to say, but we are in that category which actually benefit a lot of international

domain names technology. Emily Taylor will introduce of the findings of the report.

We have been working together and I think it's a great piece of information, excellent work and if you haven't had a chance to look at those reports, they really provide in depth analysis on the development, international names development and I'm happy we will be working this year in identifying those cases where indigenous languages as well are more and more presented and used in terms of international domain names.

So I will give the floor to Emily. I know she will be showing us the report, to I hope the Power Point could be screened, you can see it already, so I give you microphone to Emily.

>> EMILY TAYLOR: Thank you very much for that introduction and also for the partnership from UNESCO in supporting the World Report on international domain names. Since 2011 we have been working together. A couple of words about what Internationalized Domain Names are we are all familiar already with domain names used for web addresses and emails and lots of other context online. And the most common ones are in a very limited character set, A to Z, 0 to 9 and the hyphen, A to Z in Latin script, no accents, no diacritics and no other scripts. Domain names have been available for more than a decade in many scripts, whether it's Hann, Catagona, Cyrillic or other scripts, they are all out there. The trouble is that they don't have an awful lot of adoption.

So we call these Internationalized Domain Names, perhaps, they would more accurately be called Local Domain Names because they would be much more appropriate to local markets and local communities. Now the registry for the dot EU domain and as a creation of the European Commission, it's central to its role is the promotion of linguistic diversity. EurID as a company presents the languages spoken in the European Union in its customers and support. It has launched the dot EU equivalent in Cyrillic script to support the Bulgarian language community. And it's working hard to get the designation in Greek through the ICANN processes. So that would represent the four official languages in the European Union.

So if I can go to the element slide, I thought I would introduce what an Internationalized Domain Name is and why EurID has been partnering with UNESCO. It's great to see a colleague from Verisin and the Top Level Domain managers from across the world. Those associations also support our research. So looking at last year's findings which are data to December 2017, we see that there are a total of 7.5 million of these Internationalized Domain Names registered throughout the world.

And they are registered in what we in the industry would

call the generic space, so dot com and so on, that's in the light blue bit at the top chart, but the majority as you see are registered in country codes, things like dot FR, dot CH, dot CN for China and so on. That is the vast majority of country code IDNs. That's where they are.

But sadly, the number of Internationalized Domain Names has actually dropped in the last year of study. However, the consistent finding is that 2% of the world's domains, only 2% are IDNs. So if you go to the next slide, please. However, when we get over the fact that these are still quite a minority feature of the domain name space, there is a very rich array of languages and a lot of evidence that these are being used very locally to support languages.

So if you draw a map as we have done which shows the distribution of the different Internationalized Domain Name scripts throughout the world, we see that they very closely map and reflect the languages, the scripts that are demanded by the languages spoken in those countries, territories and regions. So, for example, in the Gulf of North Africa, that's where you tend to see big script domain names.

In Russia and other Cyrillic script countries that is where you see Cyrillic script IDNs and almost nowhere else. They are very localized. It is only in countries which either have strong registrar bases or where you have very mixed populations and different languages spoken that you see a mixture for the most part the script of the IDN follows the language spoken.

There is also, when you drill down, so as part of the research team at Oxford Information Labs, we corral all of the open zone files in the generic space plus the dot EU zone. So we look at those domain names and we look at the website associated with the Internationalized Domain Names. We find this isn't surprising, but it's surprising how much it is correlated that the script of the domain name signals very accurately the language that you find on the website.

So for that reason, we say that Internationalized Domain Names do promote linguistic diversity on line. Next slide, please. Even further than just looking from the domain name down to the website, if you go and look at the website language and go back up and look at the script of the domain name, you see interesting things happening as well. Now, I think Professor Mariani you mentioned how dominant the English language is on line, and in fact in more than 50% of websites are still in the English language.

And when you consider the wide number of languages spoken off line, that is a bit of an indictment in how well the technological space is currently supporting linguistic diversity. When we look instead at Internationalized Domain Names we start

to see different patterns. Now, still there is very much that sort of scenario that you were describing, Professor Mariani that the big languages win out, but there is much more linguistic diversity amongst IDN websites than the standard pattern in their counterparts.

So English goes back down closer to the level that it is respond off line, about 11% of the websites of IDNs are in English language and Chinese is much higher and so is German, so is Russian.

Looking forward to this year's support, we will be supporting the international year of indigenous languages and we will be looking closer to the other end of the scale that is usually other in most charts. What is in the other? So I drew up a chart here on the right-hand side of the screen that sort of unpacks the languages that are not represented in large numbers or high proportions. They are not going to win on the percentages, but look at how many languages there are.

There are over 80 languages in there that are signaled by those Internationalized Domain Names, including four from the UNESCO lists of endangered languages that other speakers have mentioned, Irish, Maori, Welsh and Bosque. Those use the Latin script, but we have many languages in there that I hadn't labeled individually. We have African languages, we have many, many others. That all contributes to the story that Internationalized Domain Names can really contribute to linguistic diversity on line and help to improve it.

Next slide. However, it's not all good news, and I think there was probably a hint of that in the first chart that I showed. There are probably nearly 400 million domain names registered worldwide and yet there are only 7.1 million IDNs. Why is it that they haven't shown great uptake, and without going into the technical reasons, they don't work very well.

There is not much point in having one if you can't send and receive email in a predictable way, if you can't use it for, for a website where users will always see the domain name as it was intended to be displayed to a human. Eirik talked about the importance of names to people not just in technical systems but domain names also carry that emotional quality. Names are important. They matter to people. And it's very important that they are displayed in a predictable way.

So there is a lot of work going on within the technical and language communities to try to alter this narrative and to try to make sure that Internationalized Domain Names work. There has been a lot of advances in the way that they work in websites and there is hard work at play in making sure they are also working predictably in email. There is some way to go yet.

So in conclusion, and thank you for your patience while I

show my slides. We have found over a consistent period that IDNs help to enhance linguistic diversity in cyberspace. They have great potential to change the story that we have been hearing from other speakers. They are accurate predictors of the languages that you find on web content. Minority and indigenous languages are represented in the internationalized domain space and we are looking forward to doing a bit more research into that, but the sad face at the end is there is still technical work to do, and because this is a supply chain with multiple actors and all of these services have many, many different players, all of them need to gain a sense of urgency and a commitment to work together to ensure that the technology is really up to it and serving the majority of the world's population. So I will leave it there. Thank you.

>> MODERATOR: Thank you very much, Emily, for this very in depth presentation what the situation is with IDNs. And I have one more question which I was thinking should I ask this or not because we have seen probably some of you recently the Articles in the main media or Internet media where there are some disparities about domain names and we can see domain names not only indicate language but as well they make their connection to cultural diversity and geographical names, territories, and this is what development seems to be growing.

And we have one case which is being discussed between big private companies and as well several states, for instance, domain name used Amazon. Maybe you could comment what are the, what are the possible solutions of this case because Amazon is, of course, very well-known international company, but at the same time, when we go to school, we will learn quickly that this is as well a territory which is across different countries. So how this disparate question will be solved out in Internet because it represents cultural diversity of those regions, geographical biodiversity aspect and how it could be resolved. Whether do you think? What would be the right way to move ahead?

>> EMILY TAYLOR: Thank you very much for that. So a few years ago, ICANN, which is the body is that coordinates domain name policy internationally opened an application process for new Top Level Domains as they are called in the jargon which is what happens on the other side of the dot, like dot com. And there were nearly 2,000 applications. Over several years before that, the ICANN community worked a set of rules and regulations that applicants would have to satisfy to get through. And in fact, they did think about geographical terms like city names, and country names, and had a very sort of neat and simple way of dealing with it, which is to say if you have, for example, if you are applying Bio or dot Flanders which were both applied for, you have to show that the relevant public authority or country state

does not object at the minimum and actually supports it at the best.

Unfortunately, and it's always easy to be wise with hindsight, but unfortunately there wasn't provision for those regions which don't have a -- regions which don't have a single state involved or like the Sahara, say, or in fact Amazon, where there are many, it goes through many states, but is also a very, an emblematic cultural and geographical symbol. This case has raised a lot of strong feelings on both sides. I have spoken to colleagues who work in the industries and to colleagues from the region as well, and everybody thinks it should be resolved in a different way.

Unfortunately, it's easier to sort these things out with hindsight when you know what is going to happen than it is to predict a way forward. I believe that the company Amazon is in dialogue with the countries affected of the Amazon region, and I'm hopeful that something that is equitable and something that satisfies all of the parties will be resolved in that way. Thank you.

>> MODERATOR: Thank you very much, Emily. Now, I would like to turn to my last speaker, who is the youngest among us here on the podium, which is as well is a language activist, so Rebecka Forsgren is a member of the community and an indigenous fellow and works with World Intellectual Property Organization.

But my first question would be addressed as indigenous fellow, as a language activist, what do you think, what are the effective tools to promote indigenous languages as an activist? Maybe you could share your experience what five of use, what are the most effective tools.

>> REBECKA FORSGREN: Good afternoon, everybody, and thank you for that. Like we said, I am here as the indigenous fellow for the traditional knowledge division at WIPO. And we are now launching a campaign where we promote indigenous or anyone who can translate our publications into indigenous languages.

We do this, like this is possible through something that we adopted in 2016 which is called the open access policy, and I will talk to you a little bit more about that. First, I will show you one of our publications that came into 2016 or 2017, so it's a movie. That has been produced by WIPO, and I will show 2.5 minutes of this. You it's five minutes, you can find it on the website. I don't know who I am talking to, but, so 2.5 minutes of the movie and I will tell you a little bit more about the open access policy. To promote indigenous translations of our publications. Thank you.

>> Tucked away in a tropical rain forest among the tallest trees and the smallest rocks lives the Yakuanioia people. Like many indigenous peoples they have developed their own unique set

of skills, practices and innovations, carefully passed down from generation to generation this traditional knowledge forms an invaluable part of their identity as well as an economic and cultural asset for their entire community. One Yakuanoia innovation is particularly celebrated. Once a year in late spring they come together to begin carefully harvesting the vibrant green chutes are separated from leaves and roots and placed in large clay pots. Under constant care the mixture transforms into an oily brown paste and they have a new batch of their skin marvel that can be used to heal by reasons or skin conditions like eczema or can be used as an effective moisturizer keeping skin youthful. Highly prized the paste has been bartered and traded in neighboring villages and so for many years the paste has formed the basis of the Yakuanoia community's economic livelihood.

Word of the paste traveled beyond the community, always on the search of new and innovative resources, a cosmetics company researched the active ingredients. Impressed with its antibacterial benefits they began to mass produce a antiaging face cream for distribution worldwide. Excited by its potential. The company quickly apply for a patent to protect its product and registered a trademark using the name Yakuanoia. I had no idea that the traditional knowledge of my Yakuanoia people and a Surain chute would have so much potential.

Now in the city I call home I noticed advertisements everywhere for this new miracle face cream. Having left my community years ago to study law, I couldn't help but wonder, what does intellectual property have to do with traditional knowledge? Can and should intellectual property law protect it? The company has made an investment in their product, so what would be fair in a case like this?

In search of a trusted source of information, my first step led me to the World Intellectual Property Organization. (End of video)

You can pause here. So what is this language that it's subtitles into. It's subtitles into northern Sami. So this was the world premier for this sub site will version. This is part of our campaign to spread, WIPO wants to spread knowledge about the intellectual property system, about how they can use it, and we want to make our, we want to make our publications, our materials accessible as possible and available as possible. That's why we have our publications on the website. That's why we make them available in that way, but how do we make them accessible for indigenous peoples?

To be able to make them accessible, you have to understand what is in the material, so a way of doing this is translating it into indigenous languages. So the open access policy that I

mentioned before is something that is a WIPO got in 2016 and it makes, it makes it easier and more freely used all of our publications. What the open access policy does is make, give permission to anyone to reproduce, distribute, adopt, translate and perform the content of our publications for any purpose provided that they acknowledge WIPO as the source.

So anyone is allowed to use WIPO materials for free, and without needing to get any sort of approval from WIPO first. So with this policy stakeholders, and this can be states, Government, but also indigenous peoples, so stakeholders are now encouraged to translate and reuse and adapt our publications, and they have the freedom to adapt contents for their own audiences and one way of doing this is through translating the materials.

So we want our material to be of practical use, and especially for IPLCs and that is indigenous peoples and local communities. This movie, I must say, is very practical especially for me. I think a lot of indigenous communities, especially the Sami community, we learn by doing or like we learn by seeing things more than being, than reading. So this is how I'm brought up.

So when I was coming into WIPO, I was reading a lot about like what is intellectual property, what is genetic resources, which is something we work with, what is traditional knowledge? And when I saw this movie, it was another way much easier for me to understand the content of the material, to be told like this is a very good way of using modern technology to get information out.

But then I showed this to a relative and she said, well, I don't understand anything because she don't -- it's pretty advanced English in this movie. So with this translation that we have had from the Sami Council, then much more Samis will be able to engage in these questions and we will raise interest much easier. So, yes, what we do with this campaign then is we encourage everyone to translate our publications through the open access policy. This is possible to do for free and without asking permission, and also the traditional knowledge division offers help and guidance to do this.

So if anyone here is interested in translating this movie, then you will contact us on the traditional knowledge division, send us an email, you will get the script for this movie sent to you, fill it in, send it right back, and with a little help of our communications division, then you will have your own version in your own indigenous language.

You can use that version to create an interest for this question, because intellectual property issues are a growing issue for indigenous peoples. So, yes. So this is one way of doing it.

>> MODERATOR: Thank you very much for this intervention. I had several questions about intellectual property to ask, and I think you just really answered them yourself, and I found it really as well a good way to introduce very complex issues like intellectual property issues, even if you know as you say advanced English or standard English like some native speakers, I think it's always important and as well a good practice to look at them more seriously and understand them better in your mother language because intellectual property elements, issues are very very complex, and we really can't when you share knowledge, we have many cases around the world where traditional knowledge was misused, not only what it was documented but as well misused by different stakeholders, so this is why indigenous communities, those which are involved in knowledge documentation, knowledge utilization processes, sharing with traditional knowledge what we would be aware about existing intellectual property mechanism and the opportunities what we have in place.

So knowing what we only have a few minutes, I would like to share with you during international year of indigenous languages we are running a social media campaign and we invite you to use our website which is acronym of international year 2019.org and join us on Twitter, Facebook, and InstaGram, and share your views.

Before I close the panel, for the social media campaign maybe I would invite the speakers to summarize what are the key messages we would like to deliver to the audience and as well those following us through social media.

>> EMILY TAYLOR: Thank you very much. Just perhaps just to end with a hopeful message and one of contrast, in my presentation I talked about domain names which are, you know, very much at the infrastructure layer of the Internet. But there are also some really hopeful stories, the way that technology has been used and harnessed that enthusiasm and the work and the knowledge of indigenous language communities.

One example I was reading an Article about Facebook the other day, and it said that the key thing that transformed that company from a struggling startup, believe it or not, there was a time when it was a struggling startup, was when it started to translate it's interface into other languages. And that happened almost by accident. An executive who was Spanish speaker did this and it was a wild success, and now over 80% of Europeans access Facebook in a language that is not English.

And over 80% of Facebook users do the same. Their communities project effectively crowd sources that knowledge from local communities, and puts indigenous languages into the platform, not in the content, but in the menu system so that it is an entirely end-to-end local language experience. And just

following up on a remark of Professor Mariani the impact of search supporting a language can be incredible.

There is a great paper by Sebastian Crast about a Sub Saharan African language which was then supported in search and it led to a very significant uptake into computer usage, employment, and also impacted on incomes. So there is a real story that supporting local languages is not just a good thing to do because it's the right thing to do, but it can also make business sense. It can transform the prospects of a company or and even bring those languages into a new environment where they are cool and they are revitalized in by being used in a way that is very much of this time. So thank you.

>> MODERATOR: Thank you very much. It means what multilingual makes you rich and gives you growth.

>> EIRIK LARSEN: For summary, I would like to stress that we indigenous people, we want to be part of the revolution if we can say it, but at the same time, developing new software and solutions, it has to happen with the participation of the indigenous peoples, and in in accordance with our right to self-determination. I was briefly explaining what we have done in, with some of the Sami languages, but when it comes to voice to text or text to voice, we haven't done so much and with this Artificial Intelligence and the deep learning and computers learning themselves, and improving their language, I have heard that Skype provides real time translations from English to French and back again.

And the reason why the computer has improved is that there is a large text base for computers to use. And for smaller languages who doesn't have that large text basis. So when developing new applications, we fear that our languages will be left behind again due to the lack of text, in the text basis.

>> JOSEPH MARIANI: So on my side, what I would like to summarize is the fact that I think Government should be aware that language interfaces are very important for many different reasons. I would like also to stress the fact that there is still a large effort to be devoted if we want to have proper coverage of languages all over the world, and this is really a challenge for us, a scientific challenge as well. In my opinion, multilingualism needs linguistic technologies. It will be only possible if language technology are available, and it will not be possible without language technologies, and what I hope is to discuss this further at the language technology for all conference we may have this year.

>> REBECKA FORSGREN: It has been mentioned before that not all indigenous peoples have access to computers, and it's expensive to get publications in writing and school books and such. What I want to send with you from WIPO is the fact that we

have a lot of our publications or most of our publications available on our website and technology can be a threat that has also been mentioned like sacred rituals can be filmed now very easily since we have Smart Phones and they can be spread out and such.

But technology can also be a strength for indigenous peoples, especially in the connection to intellectual property rights. So what I'm saying is use, you can always use intellectual property rights and technology for your, like to your benefit as an indigenous peoples, but you also have to understand the issues, the issues of intellectual property in one way of really getting this knowledge is to use the possibility now with open access policy to translate this movie and to translate our publications and really use technology and use intellectual property as a strength. That's what I want to say.

>> MODERATOR: Thank you very much for this excellent summary. So just I want to repeat, I think it's very good for a moderator to have kind of conclusive remarks when everything has been said so I have a few key words that I would like us to stress before we leave. If we promote a multilingualism if we support indigenous languages it means we discover new customers. It's good for our businesses, it's good for our societies. That is the first message I would like to bring out of this room.

Secondly, there is no development without participation. Indigenous peoples have to be part of the development processes. We have Asian transitions, access ancient knowledge which is little known around the world and could contribute to development. We need to be more aware what language technology is in place at our disposal, but still a lot of effort still to be taken in order to exploit full potential and integrate them at the benefits of indigenous peoples.

And last thing that I would like to share, knowledge sharing is important. So let's continue being open, inclusive, as well as accessible sharing our knowledge and I thank you very much for your patience, for your participation, for your attention, and let's continue working and discussing together in this area. So thank you very much, panel speakers, for your participation, for your availability, for your dedication and inspiration and we will be continuing working together. Thank you very much!

(Applause).

(Concluded at 1631).

This text, document, or file is based on live transcription. Communication Access Realtime Translation (CART), captioning, and/or live transcription are provided in order to facilitate communication accessibility and may not be a totally verbatim

record of the proceedings. This text, document, or file is not to be distributed or used in any way that may violate copyright law.
