

<http://intlnet.org>

Multi-Lingual Forum (MLTF)

INTLNET

23 rue Saint Honoré - 78000 Versailles - France
Tel. (33.1) 39 50 05 10
info@intlnet.org

OPEN PUBLIC BUSINESS COOPETITION PROPOSAL

Dear Peter, Mark, Vint, Addison, Debbie, Gérard, Mary Lou, Paul, Olaf, Russ, and others,

In recent years I have tried to ensure that the IETF decentralized network options, lack of presentation layer, or confused propositions would not immediately harm the distributed nature of our architectural and cultural future, in turn giving us time to organize the Multilingual Internet and consider how to simulate or integrate a presentation layer.

I have now, through you, an open public business coopetition proposal to Microsoft, Google, Yahoo!, Unicode, MLDC/Geolang, ISO, ICANN, IAB, IETF, and every concerned business and organization's decision maker to start using that spared time in the best way.

Our context

ICANN Bylaws, Section 2:

"6. Introducing and promoting **competition in the registration of domain names** where practicable and beneficial in the public interest."

WSIS Action Plan, C8. 23. o :

"o) Governments, through public/private partnerships, should promote technologies and R&D programmes in such areas as translation, iconographies, voice-assisted services, and the development of the necessary hardware and a variety of software models, including proprietary, **open source software and free software**, such as standard character sets, **language codes**, electronic dictionaries, terminology and **thesauri, multilingual search engines**, machine translation tools, **internationalized domain names**, content referencing, as well as general and application software."

Every Internet service provider and user share the same need for the normative stability that ISO 3166 provides to the WDE (world digital ecosystem) for 30 years next year, yet some of you are still uncertain about ISO 3166 linguistic support, and if it is already or still to be documented? And how do RFC 4646 and your corporate practices adapt to it?

I am pioneering a Multilingual International Network vision. This calls for a structured governance project such as the one I have engaged. It includes several specialised successive actions and fora, but the first step is the open use Multi-lingual Technical Forum (MLTF). Its role is to document the multilingualisation technical and semantic aspects, and the network metastructure layer set.

You are the main Unicode and e-Business World stakeholders or standardizers. Necessarily, our users will access and interface your machines or the one you standardied in their own language.

The need

I, therefore, need to understand where/how you want the MLTF propositions to remain interoperable with your practices. I need to know this before the ISO TC37/TC46 JAC, LTRU, IDNA, ICANN, etc. positions (which

you are currently discussing on separated or mixed mailing lists), become too technically and politically complex and confuse for anything operational and global to result from them, before a long or a very long period.

Please note that you could elude this proposal: possible inconsequences would be your responsibility. However, you should realize I am just the first one to ask about this. Even if our Multilingual propositions were not followed, other similar ones would improve and renew them. This is because they are open use, follow the WSIS action plan, match the WSIS recommendations, are RFC 4646 conformant. In our case, but I suppose others would also do, it will include a semantic addressing and addressing root file (under second level testing) and an operational simulated presentation layer.

There is also the question of the technical conditions of the **competition between several mandated lists** to solve, before ICANN can respect its Bylaws. This should benefit community trustees and users. This certainly calls for new usage architecture to be experimented with, **within the ICP-3 framework**. (we all realise that ICP-3 has prevented international open roots in permitting a fair testing. If it was not respected, what ethic would oppose private commercial projects?

This is why, in order to preserve interoperability and foster competition, we are making every effort to encapsulate all your options in our metastructure and consider them in our R&D. However, this is becoming too complex for us to follow. There is a need to understand and be confirmed of your standardier's and provider's strategy (this is also true from trade and industrial consortia other than Unicode, or SSDOs than ISO and IETF).

Your core interoperability solution

The June 8, 2007 ISO/ICANN/INTLNET meeting (<http://mltf.org/070608-adhoc-intlnet.pdf>) showed us an ISO TC37/TC46, ICANN, and (probably) IETF consensus: to remain interoperable among themselves through a practical solution. This solution seems to be a common data collection and maintenance by Geolang Ltd., advised by the MLDC, a non-profit entity which gathers some of the TC37 and Unicode tenors.

As far as I understand, the Geolang targets are:

- an ISO 639-6 as the core table of "ISO 639 as a Database", ensuring compatibility with Geolang and ISO 11179 conformant foreign matters.
- ISO 3166 (sub)parts understood as a de-facto ISO-639-6-1 (ASCII), ISO-639-6-4 (non-ASCII Geolang generated/managed ISO 3166-1 equivalent).

It seems that Geolang would mainly speak about internationalization with ISO, language identification with IETF, and localization with Unicode, in turn retaining control of the IDNs list that ICANN will use to "mandate" linguistic ccTLDs along still undefined rules and rates under discussion, in particular, during June 26 San Jose ccNSO/GAC meeting. However, we all know that these kinds of things and their financial evolution [even related to other root services, such as DNSSEC and IPv6] can take years.

I must be clear here: I disapprove of this solution; because I do not think it has a credible enough business plan. Most of all it does not address the needs of the Multilingual Internet, which obliges us to carry out normative work that is not our user's job.

- However, I must acknowledge that it is consistent with the Legacy Internet English globalization needs (internationalization, language identification, localization) resulting from the WSIS/Tunis USG obligations, as they can be perceived by status-quo oriented stakeholders.
- As such it may positively help, on a short term, industries and interests from the US/GB/AU English axis and some English speaking national communities. However, it could only be a enormous mid-term disadvantage if was irreversibly amalgated into the Internet architecture in order to, or in the hope of, "shaping the world".

This is your business decision anyway, which I can only respect and support.

Coopetition proposal

I reminded Debbie Garside on June 8 that our Multilingual Internet support system, which she partly knows is obviously ready to support your/her English Globalization system, as well as the global system of any other language entity (30,000 we understand from her published figures) to be documented in ISO 639-6.

In order to keep costs, efforts, and any possible technical conflicts as low as possible for all, I suggest that we engage into coopetition on this, in mutual respect of our members/collaborators and of their private, public, and common interests, image, and reputation.

This could result in a multipartite Internationalized/Multilingual Internet MoU. It would start with a description of what the parties accept under those terms, and the way we want to achieve a seamless direct or indirect interoperability (when this is practically feasible), a calendar for the common progression of language coverage, and a technical conflict resolution procedure. This should be enabled by a common mailing list to all the parties of the MoU.

Part of this endeavor could be a multilingual semantic addressing and routing concertation meeting and some priorities we have and I describe in annex.

Conclusion

Semantic processing, addressing, and access routing are areas where national and lingual community industries, interests, and cultural exceptions will actively coopete within, since it leads cultural artifacts from the very human imaginary to become computer and network assisted.

The market extent is so wide that I think we have no other competition than the delays that we ourselves can create (mainly for reasons explored by IAB's RFC 3869, I suppose, that I unfortunately had to correct in the RFC 4646 key case).

This experience, the current debates, the urgency for multilingual DNS view, semantic addressing and routing as part of the digital and multilingual convergence, give us a short time window and opportunity, together, to build the tools of the Multilingual Internet.

Sincerely yours,

JFC Morfin
MLTF, Moderator.

Annex

POINTS OF CLARIFICATION

To help a technical dialog, I would like to review a few MLTF points that need to be understood thoroughly.

1. Calendar

In MLTF opinion, most of the problems that we face result from the delayed availability of ISO 639 as a consolidated system. This system is foreseen as the "ISO 639 as a DB" project. The problem is not so much its date of availability but rather the total and proven incertitude of its calendar (and purpose) since it was introduced (http://www.computing.surrey.ac.uk/grid/fingrid/papers_files/Reports/7.pdf) by David Dalby, Lee Gillam, Christopher Cox, and Debbie Garside.

Your common Geolang inteoperable project is necessarily built on ISO 639-6 as the ultimate registration place. It was supposed to be online for a test in December of 2005, then at the end of 2006, yet we are now told approx. 2008. During the June 8 meeting we heard of a timetable of six months to publish a text and then a new language to be supported every six months.

These delays are quite realistic, but parallel work goes faster than sequential work (yet it also calls for more man-power from a limited existing pool of experts). The wikipedia suggestion is certainly an interesting one, but I have some difficulty integrating it in an ISO 11179 process without some common thinking and further standardizing effort.

This is why we need to have a clear and transparent project based on a technically clarified and published NWIP with a very documented project business plan, committed credible dates and financing. This is because we all know the delays that are incurred by delayed payments to small structures. From a contrary position, a small structure is certainly a factor of speed and flexibility.

2. The lingualization, globalization, and multilingualization terminologies

We can dispute for years on the rationale or even the existence of these terms instead of mutually translating them. They are commonly used terms, in debates, the press, WSIS documents, and IGF preoccupations.

The MLTF accepts:

- "**lingualization**" as the use of a language in a process, such as the Internet standard process.
- "**plurilingualization**" as the optional support of several other languages by the process language.
- "**multilingualization**" as the technical possibility to support every language at the same level of usage.

In Unicode terms, we understand that "globalization" is a plurilingualizational process. From their own inputs, **language globalization** = **internationalization** of the medium + **identification** of the foreign language + appropriate **localization** of the ends.

Using that terminology, **multilingualization** = **globalization localization**.

MLTF accepts "**polynym**" to qualify the multilingual name list of a language invariant concept.

MLTF accepts an "**ontology**" as a collection of agreements upon a vocabulary of terms and meanings.

3. What is a language in the ISO 639 and in the ISO 3166 perspective?

During the 6/8/07 meeting, I felt that my ISO, ICANN, AFNOR, BSI, and Geolang interlocutors confused the four metadata definitions of a "language name" (ISO 639), "language" (ISO 639-3/-4 and IETF), "administrative language" (ISO 3166-1), and "linguistic community". MLTF understands:

- a "**language**" as a human/machine relational protocol that can be differentiated in reference or through an algorithm from a corpus wearing the language name.

- a "**language name**" as any terminology that is used to designate a language.

- an "**administrative language**" as a communication tool, in which the specification of the usage is not a technical restriction to trade in the considered territory (as per WTO rules).

- a "**language community**" as all of the persons and processors, possibly from a given place, country, or continent, using that language as a common relational protocol. We consider that when an administrative constituency is documented, it has an identification priority because States and administrative regions are the only self-defined human distributed referential.

I wish to note that the MLTF definition is totally consistent with computer languages (C, Basic, Perl, in any human script and language). This is true for other kind of existing or prospective forms of language, such as the coming neurolanguages.

4. The mode issue

ISO 639 and ISO 15924 are script oriented, RFC 4646 is script influenced, and MLTF is **mode** oriented. This means that we used to sort language information whatever the mode (script, oral, sign, etc.). It should, therefore, be understood that our use of "mode" is a superset of others' "script".

5. The library issue

Our tagging system is entirely RFC 4646 compatible as requested by the WG-LTRU and uses the "x-" private use option.

We will suggest the use of the "x-n-" as a world digital ecosystem private use extended header.

Due to the size, simplicity, and generalized vocation of these tags, it is likely that libraries supporting our tagging system will be developed, which will support our system as an extension. Our intent is to publish all information in order to permit any library and user to remain interoperable at every lingual, modal, and administrative level.

6. Semantic addressing and routing

We obviously all know that we are starting to compete over what will most probably be the leading market of the 21st century: semantic processing and networking.

The experience of the DNS and of the IPv6 addressing, the fundamental problems raised by ROAP (IETF debate on these issues), and the limitations that could be imposed on the ISO solution, should lead us to investigate a joint civil, private, and public consideration of the the semantic addressing standardization. This issue is obviously fundamental to Metacomunications among natural (human brain) and mechanical (computers) semantic processors.

6.1. The MLTF, therefore, is considering hosting a concertation meeting on multilingual generalized semantic and pragmatic addressing and routing, at the end of September 2007 in Versailles, prepared by a meeting in Paris during the second week of August. At this stage, all I can say is that the meeting will happen. I certainly welcome all interest and suggestions. We have access to a prestigious 40 seat lounge (where the MLTF was incorporated), but can accomodate various meeting types - possibly with spouse programs by the Turism Office.

This meeting should be state of R&D oriented. It certainly involves at some various levels TC46, TC37, and JTC1 oriented concerns, as well as computer, software, telecommunications, and datacommunications research centers, ccTLDs, archivists, museum and libraries, and cultural endeavors, etc. as well as Governments. Its target is pragmatic: to understand at network access, routing, and addressing levels how ontologic essence, referent, context, personalisation are to be network supported.

The choosen date could permit a Semantic Addressing and Routing workshop at the Rio IGF Meeting. Rio could be the place and the time to introduce a Dynamic Coalition on the access and standardization of the "multilingual information of common good". This information package is all what is necessary for an equal

opportunity access usage of the WDE (world digital ecosystem) in every user's language. This should certainly help accessing some of the millennium objectives. It would most probably also be a good basis, test bed, and field of experience for further industrial, commercial, and open source propositions.

We could then target together a more formal international meeting, as part of the Year of Languages, in 2008, probably in Barcelona as preparation for the IGF Dehli meeting.

6.2. MLTF participates to an attempt to catalyse grassroots French and Asian considerations on the way to implement an IPv6 and or OPES based presentation layer equivalent in a transparent way to the network technology, and to test it along the ICP-3 criteria.