APPEAL TO THE IESG AGAINST AN IESG DECISION

This is an appeal against the IESG decision to approve the RFC 3066 Bis Draft on November 15th, 2005. It documents the context of the appeal. It discusses the points raised against the IESG decision and also proposes solutions to suggest the WG-LTRU.

1. Context

This part documents contextual or architectural points which this appeal does not raise. These may be raised if this appeal has to escalate to the IAB. These points help to understand the context of this appeal.

1.1. new challenges

Languages are symbolic ways to convey information through different modes (spoken, written, signs). These are the protocols of human layers exchanges. These are parameterised by referents (what is accepted by all their speakers). These are used along various personal styles and in different contexts (what is accepted by all the participants to an exchange). Languages are currently under major evolutions:

- documents have become computer assisted, with the emergence of "architexts" (smart source of the presented text), with new memorisation, presentation, multilingualisation approaches.

- new human technologies may lead to different referents for a same language. A text may have consistent orthogonal readings depending on the trade of the readers of a same language.

- languages are networked. They are used:

  - in polylogual multilocutors contexts documented by underlying or environmental meta-elements.
  - in multilingual and multicultural local competition.

Symbolic ways are out of the scope of the IAB and IETF, but protocols over a network are in their scope. However, the mutual impact of protocols, symbolic ways, modes, styles, computer, trade, networking and cultural competition make the identification and the referencing of networked languages an issue of responsibility and of competence, with technical, societal, cultural, economical, political and vernacular usage implications wider than the scope of any existing SSDO.

1.2. globalization vs. harmonisation doctrines

Globalization is understood by the IETF as internationalizing the network and localizing the end. The target is to remove the lingual barriers between a supplier and its foreign clients. The Internet, as agreed at the Tunis summit and documented in the reports of IAB and ICANN meetings on IDNs, is the internationalized US data network system. It empowers the American language, culture, economy, etc.

The task now is to permit an equal empowerment opportunity to the other 96 % of the mankind. "What ever you can do in American, you MUST be able to do it, with the same comfort, in all other 20.000 languages listed in future ISO 639-6". It calls for a, most probably rewarding, new architecture.

The architecture to address this challenge is multilingualisation. It can be a cross-globalization for every language; a universalised core and a personalisation of the ends; a multilateral architecture … In any case it should support the IETF US internationalization, possibly as a default – if it permits.

Multilingualisation will be a never ending process. It will have to keep pace with the changes in the human society, technology, and cultures. This constant harmonisation will probably be very complex. It will be a dramatic change, from the mono[lingual] Internet we are used to, to the multi[lingual] Internet we need. This evolution will not be limited to its basic requirement: brain to brain interintelligibility.

1.3. conditions of the debate

I came to the IETF to offer RFC user QA over multilingual issues at the WG-IDNA. I found no interest in the users' needs. I found a lobby (leadership?) wanting to "influence" the ways the world "design, use and manage" the Internet (RFC 3935). Network hysteresis does not permit to disregard errors engaging usage. Experiences in naming and numbering have shown these may delay the network development for decades.

Creating an opposing lobby was my first move. I quickly saw it would only be damaging to the IETF, a bigger waste of time and non productive. This is why I chose to give the necessary time to common sense to prevail, in exposing their mistakes in a way they could forced to correct some of them. The democratic method for that is work and filibustering. Filibustering is not pleasant. But it permitted to obtain what users' protection demanded:

- to get the lobby discussed above to expose itself, and its (RFC 3869) commercial sponsors
- to get done most of the work needed to clean the messy December 2004 version of the Draft.
- to get this Draft approved only after the Tunis Summit made it a US local proposition.

It did not permit yet to make it useful to users. This is the reason for this appeal.

Work allowed me to build an alternative open solution. Not to be disruptive I need an IAB guidance on its introduction. I used the appeal mechanism against the aforesaid lobby to request it. I now wait for it.

1.4. purpose and context of the Draft

The purpose of the WG-LTRU debate and (non-considered) charter is in fact to build an authoritative Language Name System (I call LNS for simplicity). This LNS is to be consistently used to name languages, locale files, lingual products, contents and services, etc. and even possible lingual TLDs.

Yet, the WG-LTRU debate has shown the intent was not to build this LNS as the basic building block of the open Multilingual Internet, a Multilingual Internet the world calls for. It was to constrain its ABNF and to use its registry management to influence designers, users and managers (RFC 3935) towards an internationalized Internet along the Unicode consortium' vision. Forging such an exclusive, through the single IANA langtag registry, reminds the ICANN exclusive over the Domain Names registry. Network stability and good governance should then lead to an IETF/Unicode/IANA MoU over Language Names issues. It can either be taken as a commercial war declaration on cultures, or a decision to balkanise the Internet.

1.5. not documented resulting obligations

RFC 3066 Bis was to welcome ISO languages and script names within the Internet standards. The Draft does not consider the extensions under way of these standards. It does not discuss what a structural analysis shows as a layer violations(wrong location of the Unicode table, taken as a super charset instead of a registry). The mere allusion to such a possibility was considered as … insulting.

- ISO 639-4 and –6 may partly or totally obsolete RFC 3066 Bis. Long term JTC1/SG32./W2 efforts, ISO 639-6 and ISO 11179 conformance may lead to much more powerful propositions, giving a corrected RFC 3066 Bis its place in the global language information jigsaw.

- We all need, and will obviously implement, an open compliance with ISO 11179 and with any other SSDO standard. WG-LTRU decided against ISO 11179 conformance, bridges and compatibility.

- The support of the Unicode OS unification CLDR locale files project is one of the targets of the WG-LTRU. Its discussion was denied. It is promoted since the Draft has been approved.

RFC 3066 Bis makes the IETF to support a unilateral US industry proposition. It imposes on the IETF (RFC 3935) competence and responsibilities demands which it cannot match. The IETF is now going:

- either to manage a not yet considered, defined, organised and budgeted compilation, verification and dissemination system, eventually much larger than the DNS, in an area foreign to its engineering skills and legitimacy.
- or to dismay the whole Internet community as it did for the IDNA, but on a much, much larger scale.

Without a complete review of the undertaken responsibility and obligations, the first option is unrealistic.

2. Contested points

The points documented in this part make the appeal against the IESG decision

2.1. non-documented management aspects

The presented Draft does not document the expected usage of the langtag registry. Without any of the concerned linguistic, cultural and political authorities being involved, there are hundreds of languages now documented by a now started IANA registry. They should soon extend to a few tens of thousands. The architecture of this registry is not intended to link other registries. For example, every XML document may at some stage call for an updated direct or indirect access to the IANA langtag registry. Should every Internet user do it only once a year, it would mean more than 30 (average) 2 Meg file downloads a second.

There will be a need of a large number of yearly updates, additions, changes (variants, new languages, changes in national structures, etc.), calling for much more user updates.

- the current "ietf-languages @ alvestrand.no" registry management mailing list has handled 72 registry entries in ten years.
- there is a reasonable expectation that developers, organisations, users, etc. start relying on the RFC 3066 Bis langtags and on its registry, and that this registry cannot scale and support their current or future legitimate extension expectations.
- there is a reasonable expectation that traffic needing langtag information to proceed be blocked in an undetermined future, due to a beyond practical possibilities and lack of access to this information.

This is not documented in the security considerations part.

2.2. Competition to the IANA

The presented Draft does not want to consider alternatives solutions in naming and documenting languages. This creates a situation of competition between the IANA and other registries. No one has been accustomed to such a situation up to now.

If such a competition (as it is likely from the current state of the art) find enough support and usage, this may lead to a balkanisation of the Internet where IANA compliant only systems would not be able to use common practice procedures partly supported by the IANA. (Due to the relative volume of data and related data the LNS may represent, it is likely that this will lead to switch of the users from the IANA to the LNS system).

This is not documented in the security considerations part.

2.3. ABNF confusion confirmed by the IESG

RFC 3066 Bis uses a very restrictive ABNF. This format uses one character or figure prefix, followed by a dash, to introduce one or several specialised subtags of 8 characters or figures. There is no initial

prefix if the first subtag is a language code. Most of the 36 characters and figures are not currently assigned a meaning as a prefix. The "x-"prefix opens a private use area. I represented several times that external SSDOs and private users needed a larger space than 8 characters of figures. It has been repeatedly answered that users could use several 8 characters or figure strings to fully document what they wanted.

"en-x-ietf2006-ustx7071" is a valid langtag. This format is dissuasive. Its documentation is confused. An example of progressive langtag abbreviation documented that it obeyed to the common rules. The AD, confirmed by the IESG, decided it was a typo, removing the only clarification on the private use format. Harald Alvestrand documented that this format is dangerous. There is no other way than walled gardens to prevent two groups of users to legitimately use the same strings with a different meaning. Everyone can imagine the dangers this may represent.

This is not documented in the security considerations part.

2.4. non-documented dangerous types of utilisation

The external observable major change from RFC 3066 to RFC 3066 Bis is a distinctive format. The WG-LTRU's Charter demanded it for "easily identifying the role of each subtag in the language tag, so that, for example, whenever a script code or country code is present in the tag it can be extracted, even without access to a current version of the registry".

This leads to a dramatically easier use of IETF langtags in order to:

- profile traffic for cultural, national, racial, religious evaluation
- cultural, national, racial, religious searches in search engines
- cultural, national, racial, religious relational "marking" in using retro-meta-spam: one interlocutor "marks" his traffic (web pages, mails, etc.) with langtags. The returning traffic designates the persons who were able to understand it. These persons ignored they were victims of a meta-spam.
- filter national traffic for specific content, as against human rights or democratic behaviour

In addition, it is to be noted that:

- the information made public through langtags meta-spam on public mails and sites permits to collect information on private persons violating privacy laws of many countries.
- this hidden privacy violation may endanger the business of search engines which may start offering information (ex.: please search Google for "ar-arab-us") which may be illegal in some countries.
- this may become criminal activity in some civilised countries when that information concerns cultural, national, racial or religious information on the persons.
- the documents such as RFC 3066 Bis which permits such a crime and incites Internet users to implement solutions permitting such a crime may be deemed criminal themselves in some countries as inciting to racial hatred. This engages the responsibility and the credibility of the IETF, in particular of its leaders and of the IESG. This would certainly be detrimental to the whole Internet.

This is not documented in the security considerations part.

2.5. lack of verification

No validation tool is associated to the RFC 3066 Bis langtag proposition. It is entirely based on trust with no security-associated feature. The langtag of a text can be untrue. There is no documented algorithm to permit verifications (for example the number of occurrence of " the " to confirm the claim of "en-" langtag.

Actually, the debate on the further filtering document shows an opposition between its author and Harald Alvestrand over the encouragement to use the langtags not to document the text of the content, but to drive the filtering system. Such a practice will defeat the whole purpose of the described best "practice".

This is not documented in the security considerations

3. suggested solutions

This part documents the suggestions that IESG should give to the WG-LTRU to consider in order to tackle the issues raised above.

3.1. update of the security considerations

The first proposition I can suggest is to rewrite the security considerations. This was denied to me by the WG-LTRU, when I started considering some of the points above. The current text is only a copy of the RFC 3066 considerations, as if the RFC 3066 Bis did not introduce any new risk.

3.2. management of the resulting obligations

The unilateral centralisation of the IANA functions makes them a political/commercial target. It also confines the IANA to a limited number of issues. In addition, the RFC 3066 Bis now affirms the internationalization approach. The NTIA publicly confirms that the IANA could be sold to the private sector (names have been rumoured). This puts the IANA for the first time in (commercial) competition with the work engaged by other SSDOs or grassroots processes and their own clearing houses.

I engaged such a process. I am here to represent the rights and interests of such endeavours which the WG-LTRU disregarded, as an IETF DoS. My DRS project (distributed registry system) is an extended and distributed approach of the IANA. A LNS (language name system) is mandatory to such a project. I will support RFC 3066 Bis langtags, among all the others system we may encounter.

The multilateral management and distribution of an inclusive LNS proposition, in a way comparable to the DNS, is probably the only way to get a stable, open, flexible, democratic, and scalable solution. This is what the WG-LTRU should document. I am certainly not alone in considering such architecture. But I may be the only one desiring to maintain consistency with the IETF propositions.

3.3. equal lingual opportunity position

I suggest the WG-LTRU should be asked to include in the Draft, the following equal lingual opportunity proposition.

"The purpose of technology is not to ensure that everyone should have an equal opportunity regardless of his language, its purpose is to allow for this goal.

"In its domain, this document wants to allow everyone to freely share the cultural life of the computer and Internet community, to enjoy using its solutions on an equal linguistic, cultural, technical, economical and commercial basis. It also aims to allow everyone to share into the benefits of the technology and its advancements without distinction of any kind - such as race, colour, sex, language, religion, political or other opinion, national or social origin, property, birth or other status and education. Therefore the goal is to build a common technical standard for all people and all nations to be used, or embedded in any technical solution, without any limitation resulting from the script, the language, the references or the context of its utilisation environment.

"Taking into account the granular nature and the diversity of the world's digital ecosystem, as well as the requirements of its technical convergence, the authors of this technical memorandum strongly advocate its free, secure and stable implementation to serve the rights and freedoms of its users at a global, multinational, international, cultural and lingual, national, local, and personal level, in the respect of sovereign laws and jurisdiction of each State, of the empowerment of local cultures, and of its intergovernance by subsidiarity, in each of its government or private, community or individual application."

3.4. addressing the users need

The concept of indicating the language attributes of texts in the documents themselves is a totally new concept, only permitted by the use of architexts through historically very young, blossoming and often non-compatible formats. RFC 3066 Bis (due to the importance of the Internet in content transport, access and archiving) wants to make this revolution in human practices simultaneously:

- de facto mandatory and global
- ruled by its sole ASCII ABNF and ISO selected lists.
- documented by its sole IANA Unicode based and ruled registry.

It also wants to make that change as a BCP (description of existing practices) while,

- it did not consider its consequences in the legal, privacy, cultural, meta-spam, etc. areas.
- it did not use other names than those selected by Unicode for leading mostly written languages.
- it was not approved by any concerned legitimate cultural, societal and political representation.

Should the US law not object its consequences; such a proposition can be standardised for the Tunis accepted internationalized US Internet. It could even be accepted as a working global default proposition for the convergence of the world digital ecosystem. But it MUST support other language naming systems which are able to replace it along with the evolution of the state of the arts, the user common practices, the legal obligations, the international agreements and the not yet considered security aspects.

The diversity supported by the 'tag' URI scheme (RFC 4151) provides an off-the-shelves solution which matches these criteria.

Some among all the naming information provided by the langtags can also be obtained from a URI (or IRI). There are three ways of using RFC 4151:

- to use it as a replacement of RFC 3066 Bis. This would be a final balkanisation of the Internet.
- to introduce a one-letter/figure+dash sequence as an escape sequence to RFC 4151. This however breaks the langtag concept and continuity.
- to create a users RFC 4151 ABNF conformant space. It will be introduced by a figure+dash sequence (to be script transparent). Abbreviations of this space will obey RFC 3066 Bis rules.

  My proposition is "'0-'+RFC 4151 ABNF". Others may be devised.

  This proposition was disregarded by the WG-LTRU Chair. I was banned for supporting it. An appeal confirmed that odd position.