

MULTILINGUAL INTERNET

Multilingual Distributed Referential
System

networks result from common adherence to accepted standards, parameters, addresses, and names. This is why its referential system in turn builds a network.

INTERNET MULTILINGUE

Système de Référentiels Distribués
Multilingue

les réseaux résultent d'une adhésion commune à des standards, paramètres, adresses et noms connus. C'est pourquoi son système référentiel construit un réseau.

GENERAL INTRODUCTION

In collaboration with a number of partners, INTLNET has embarked upon a project of collecting, managing and disseminating the information necessary for the deployment of a Multilingual Internet through the use of **MDRS/SRDM** (multilingual distributed referential system/système de référence distribué multilingue).

This project is based on an analysis of the relational digital ecosystem summarised hereunder.

- There are three technical layers of the digital network :
 - **interconnectivity** (lines and switches : telecoms infrastructure),
 - **interoperability** (internet superstructure, MPEG, portable telephones),
 - **interintelligibility** (metastructure of relational spaces between persons).

INTLNET uses an extended model of digital networks that develops these three main layers, including the OSI model layers, and documents the intelligence of networked shared usages (brainware concept).

:

- There are four linguistic layers:
 - **universalisation** (independence from languages, usually digital)
 - **lingualisation** (technology for a given language [English for the Internet, French for Minitel])
 - **globalization** (the extension of lingualisation to reduce barriers between other languages by internationalization [Unicode] of the environment and localization of ends)
 - **multilingualisation** (globalization of all languages). The first requirement is to identify languages in a way that is clear to all speakers, human and machine, and subsequently their spaces of exchange.
- There are five layers of human communication (and of authority management - governance) that overlap in varying degrees according to technico-social density.

- **listening** : I obey,
- **soliloquy**: I decide for myself,
- **monologue**: I decide for others,
- **dialogue**: I decide with others
- **polylogue**: I decide for myself considering what the others say and do

Current technical and multilingual convergence is a generalisation of the capacity for the global polylogue we need, have begun to experiment with, but for which we still do not have the appropriate tools (example: spam that originates from a technical inability to control the ubiquity of the system) nor the needed mechanism of intergovernance, as we still are subject to a single governance culture.

- The problem encountered by the internet is its technical, linguistic and governance-related unilateralism, a legacy of its American academic origin that has led to its development as a monolith that – owing to a lack of imagined or desired interoperability – should replace other technologies while remaining under the thumb of the IANA, and being limited by its centralisation. In the face of this rigidity, and in consideration of the resultant opposition, there are two possible schemes:
 - technical **fragmentation** (NATs, DNS instability etc), more or less rectified by ICANN's contested authoritarianism,
 - or the **decoupling** of the development of the various technical, linguistic and relational layers enabling a concerted multilateral progression
- The multilingualisation that we envisage is therefore **at the metastructure and usage layer**, independent of networking technologies (its development is common, ensuring total continuity for the internet, the media, e-learning, ecommerce, the web, mobiles etc.)

It is therefore about identifying, documenting and e-supporting the linguistic relational spaces (mailing lists, forums, readerships, trading areas, cultural spaces, economic basins, academic, residential, professional areas, etc).

- To achieve this architecture of the **networks of the network of networks**, an intelligent interfacing technology is needed between the application layer of the various network technologies (OSI layer 7) and the user application (server). This "interapplicative" layer should be compatible with every networking technology and potentially every semantic, making them content transparent. This "interapplicative" layer is explored by the middleware, and is what IAB calls the "edge".

To concretely establish the **Multilingual Internet** through usage, INTLNET is has undertaken the organisation of an **MLTF (Multilingual Technical Forum)** with four priority areas:

- The "virtual cartography" of the **relational geo/socio-linguistic spaces**

Where human and human/machine exchanges are recorded. The aim is to achieve a "langroot" registry giving access in a structured, documented, cross-referenced and named manner to the greatest number of parameters. These parameters will be

updated through their description from the largest number of sources (SIL, Linguasphere, ISO, Unicode, IETF, WikiPedia, Francophonie, etc.).

This project uses the open source work of David Dalby (<http://langtag.com>). Each space will have a site, a working group and a documentary base.

- **The path of linguistic e-empowerment**

To properly understand all the requirements that need to be supported, and to enumerate the methods that are already available through the ISO standards, the various SSDOs, the conditions of their interoperability (Eg. IETF: langtags and IDNs), the evaluation of technical possibilities and the industrial requirements for supporting the e-emergence of a language.

This is the work undertaken through contributions such as those of Marcel Diki-Kidiri.

- **Technical reflection** intended to support the availability to users of software solutions and information bases needed for a multilingual perception (the language of everyone being THEIR pivotal language in the digital world).

- **Ontological gardens:**

Conditionally interlinked ontology forests (intelligence) on the basis of ISO standards, JTC1/SC32/WG2 and reflection adapted to multilingualism and to the distributed nature of digitalised relations. This should enable all compatible computable ontology to be registered, obtained and used.

This part will probably be built on new architectural concepts based on a new paradigm of understanding information once it is distributed, freely managed by all, and multilingual (Reflection Group, France)

- **Dissemination and universal availability of data**

The data of these ontologies document the common relational system. It involves a universal common good, of which DNS is a part: its useless and even detrimental current centralisation should not be replicated.

What is needed, therefore, is: an open compilation model, management tools, reference centres, a non-proprietary distributed dissemination system, an economic model for the compatible dissemination of proprietary information, and public contributions for essential (health, security, alerts) and general interest information.

The first building block is the “langroot”, the cross-referenced, open compilation of linguistic registries enabling the linguistic organisation of the project itself (meeting on this subject planned for the beginning of October for interested parties).

- **Practical organisation and individual e-empowerment for all**

Autonomy of all in :

- Choice of one's **language**
- Choice of one's **referents** ("local" information bases common to a linguistic space [eg. 'Académie Française dictionary])
- Definition of one's personal and relational **contexts** (individual variations of referents) and authorisation for their access
- Protection of one's **manifests** (object descriptions)

requires the specification and development of tools and services. They should allow interactions that are transparent to the technologies in use and to the other digital ecosystem operational methods. In this way, separate development of each system, and stability of their interoperability will be enabled.

(Prototyping in progress).

- Documentation of the **first layer of multilingual naming**

There are various multilingual naming services

- Multilingual domain names (eg. China)
- Keywords (eg. Korea, Turkey)
- aliases (private keywords)
- left hand side of email addresses
- digital addresses

By way of a daily update, INTLNET has for several years been depicting the primary DNS zone as it really operates. It has been extended to the Multilingual Internet. In particular, it includes:

- the production of a **root reference file** including the Registries and name servers of linguistic domains.

<http://intlnet.org/intlfile.txt>
<http://intlnet.org/intlroot.txt>

- and the **stubs** needed for their use as a complement to ICANN's root file.

The use of the "punycode" function enables an ISO 10646/Hexatridecimal conversion (figures from 0 to Z and the symbols « . - + »). This function is generally used to support Top Level Domains (TLDs) primary zone suffixes and linguistic email addresses.