http://intlnet.org

INTLNET
71 rue Royale - 78000 Versailles - France
Tel. (33.1) 39 50 05 10
info@intlnet.org

ITU - UNESCO

Symposium global commun pour la promotion de l'Internet Multilingue
Joint global symposium on promoting the Multilingual Internet

introductory folder
dossier d'introduction

MDRS
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.

1. introduction to the MDRS
2. notes sur la présentation PPT

MDRS Draft can be found at:
http://intlnet.org/mdrs.pdf

SRDM
Système Référentiel Distribué 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.

1. introduction au SRDM (en anglais)
2. notes sur la présentation PPT (en anglais)

Le projet de document SRDM est sous:
http://intlnet.org/mdrs.pdf (en anglais):

MDRS

Http://intlnet.org/mdrs.pdf

a Multilingual Distributed Referential System

INTLNET
J-F C. Morfin
jefsey@intlnet.org

The MDRS project concerns the study, documentation, and deployment of multilingual open ontology based distributed referential systems. Its purpose is to empower each user, users' community, and their network relations with their own interoperable distributed evolution of the IANA functions. It consists of Relational Context Resolvers (RCRs) calling on Common Reference Centres (CRCs). They are ISO 11179 conformant registry servers.

Every IANA registry will be made accessible among the MDRS global references. The MDRS is, however, much wider in scope (number of covered topics), diversity (number of sources and freedomof their selection) and use (types of extended network services and user applications). This is why the MDRS represents a significant progress for network usage and for its governance.

The MDRS has no lingual bias and is designed from the onset to be multilingual. The Langroot Project concerns the MDRS language area and its root file.

- It understands languages as a linguistic community relational brain to brain/CPU protocols.
- It supports the recent evolutions of computable, screened, and networked languages.
- Its main registry desires to cross-references all the different existing languages codes
- Its other registries concern the documentation of attributes such as scripts, related geographical area, official language, ontologies, locale and comparable files in the personal, cultural, and societal areas, modes, styles, on-line spaces, authoritative referents, terminology sources, events, etc. etc.

The support of the Langroot Project and the review of its computable ontology, concepts description, basic content, etc. are complex issues. It calls for numerous types of expertise, in each of the supported languages. This will be carried out by an ad-hoc Multilingualisation Task Force (MLTF).

The Multilingual Internet should empower local languages. This means an equal lingual opportunity that is technically granted to everyone. This is a Human Right of high economical, technical, and political value. The Internet architecture was not designed with multilingualisation in mind: the Internet is not language (any language) transparent. Its protocols are "English inside".

A two layers strategy is possible to address this "e-bias":

1. **Globalization** as proposed by UNICODE/IETF

   - internationalizing the network environment (extending the ASCII capacities via different option)
   - localizing the foreign computer using a set of parameters documented in a "locale" file.

   It does not remove the Internet lingual bias: English still acts as a pivotal language between non-English interlocutors. Languages index concepts: if their index is a word and not a number, the language of that word has the pre-eminence. Globalization also does not scale. This comes from the complexity of a semantic core. UNICODE attempts to unify language tags, support libraries, and locale files' nomenclature. Their CLDR locale file project supports 128 languages. This leads in turn to the death of many languages and massive domain losses. One must go further.

2. A progressive **harmonisation** towards a fully Multilingual Internet architecture

   It consists in considering the English globalization as a default solution regarding what it permits for the English community as the target for every other language community. The MDRS supports this with a langroot open to 30,000 language entities.

This strategy permits to immediately specify, implement, and sell multilingual metalevel services; to support cross-technology naming services; to provide universal digital reference for language agnostic protocols. This should commit the network towards a mid-term credible multilingualisation. Such a commitment should in turn induce the necessary R&D non-commercial funding by Governments and via the non-profit sector involvement that the IAB calls for in RFC 3869.

INTLNET
71 rue Royale
78000 Versailles
France

http://intlnet.org
info@intlnet.org

INTLNET
THE MDRS LANGROOT

Empowering multilingual Internet relational spaces

J-F C. MORFIN

The networks of the network of networks

Analysis shows that the Internet is the multilateral, multitechnology, multilingal **concatenation** of a global continuity, and the **partitioning** of this continuity into a relational space diversity at each structural layer:

- infrastructure (hardware): operators networks, extranets
- superstructure (software): intranets, classes, externets (internal external network look-alike)
- metastructure (brainware): TLDs, user oriented spaces of exchanges, trust, and services (SETS)

The discussed MDRS is a metastructure layer project.

Relational Spaces

These relational spaces may result from different:

- architectures and technologies: telephone and Internet, IPv4 and IPv6, innovation transitions.
- national, cultural, lingual, local, economic, professional, community, private, or operational interests.

They usually are identified by a "root name" (ex. TLD, telephone area code). It will also be the name of their class of users, group of servers, specific protocols, registry, referential centre, and governance.

Identity, Universality, and Subsidiarity

Relational spaces must be empowered. This results from the three principles of:

- **identity**: to describe itself to itself by itself and for itself. In its own vernacular, along its own culture.
- **universality**: a global service continuity for equal access and opportunity whatever the point of access, technology, language, applications, services to the relation, etc.
- **subsidiarity**: to sovereignly govern its relational space and interoperability, in mutual respect, for risk containment, self-control, technical, societal, and economic efficiency, and to address specific situations and innovation, etc.

This means the full control of its reference centre by its community governance. In the global Internet, this reference centre is the IANA. The IANA possible sale by the USG would match these principles. Failing to address this network structural need in turn leads to balkanisation, which results from not pre-empowering naturally pre-existing relational spaces.

MDRS, an evolution of the INTLFILE

Intlnet was created with Tymnet in 1978, initially to help Interational Packet Switch Network public and private relational spaces to attain seamless empowerment through:

1 reference information: data collection, analysis, and dissemination
2 governance catalysis: support, R&D, testing, and project management
3 intergovernance assistance: consulting and secretariat services

The MDRS project is a state of the art evolution of the daily INTLFILE service it publishes on the top zone.

Ontology Forests

MDRS stands for “multilingual distributed referential system”. Its purpose is to permit each relational space to maintain an interoperable computer accessible description of its own reality and commonle chosen references (computable ontologies). The same concepts should be used everywhere for interoperability reasons, but the content will depend on the choices of the governance of each relational space.

The MDRS project is technology transparent and is a distributed registries network architecture that organises and intensively operates ISO 11179 conformant computable and interlinkable ontology “open forests”.

MDRS architecture

The MDRS architecture includes:

- a metasystem of computable and interlinkable ontology root registries and IT management tools.
- Common Referent Centres and OPES Relational Context Resolvers logics*
- a protocol and solutions to interlink computable ontologies, including the IANA
- an inter-operating system description to support open network extended services (ONES).

Three main roots will document the MDRS’ own structure, languages and networking.

Networked Language

The first need was to identify what is a networked language compared to a spoken or written language.

- major evolutions have occurred:
    - the **screened** mode: a script is no longer a way to memorise, but rather to extend the content
    - the **computable** architexts: a new complex, possibly multilingual, and multi-purpose area.
    - the **networked** mode: importance of polylogue over dialogue.
    - the **abbreviated** mode: SMS, various short-hand systems, coding.
    - etc.
- this leads to the understanding of a language as a linguistic relational space brain to brain/CPU multi-channel protocol, eventually (one or two decades?) identified by its algorithmic multimodal signature.

Globalization and Harmonisation

A multilingual system must address the “e-bias” of the Internet and computers: they are “English inside”. The users' lingual spaces are not and do not want to be. A two-layer strategy is:

1. **Globalization** to remove lingual barriers with foreign users, in internationalising (extending the charset) the Internet and localising (translating in local words and formats) its ends. It supports **interoperability**.
2. **Harmonisation** to progress toward a global multilingualisation ensuring every language the same local and global empowerment and **interinteligibility** as English globalization does for English.

The current common proposition that remains at the globalization layer does not scale, which addresses a multilateral need as an option of a unilateral solution.

Registries interoperability

Difficulties arise from the diversity of language coding systems in a limited number of languages.

- the language definition will avoid the subjective evaluations of variants, dialects, sociolects, and idiolects
- the root must offer a language codes cross-referencing matrix, expressed in every language

Universal tags (unitag) are used. They are concept and element ISO 11179 identifiers. They are made of a sub-address (database or IPv6 sub-address), cloud of lingual names and aliases, as well as a status-date-origin stamp. They can chain local or foreign registries through qualified computable interlinks.

Interoperable interlinking

The language unitags (langtags) of the current experimental langroot offer equivalence between ISO 639-1/2 and 3/5/6 Drafts, Linguasphere, new description tables, etc. and include computer languages. Interoperability must be at two levels:

- data interchange via ISO 11179 conformance and support of ISO 20944 exchanges,
  - directly from the source.
  - through an intermediary source regularly ported from the source. Ex. IANA registries.

- conceptual interoperability. The main problems found thus far:
  - confusing concepts and format descriptions.
  - deliberate concept constraints (mostly for commercial reasons). In such a case, an intermediary source must restore interoperability through an inclusive superset of its source.

Example: let us imagine a local language registry that deliberately ignores the European Union administration language variants. The intermediary source version of that registry should add the “EU” region to permit interoperability, but qualified links should remove it in the concerned local relational space.

The langroot Structure

The langroot is a key building root in the MDRS ontology forest. Its langtags fluidly (open framework) describe:

- classes: language codes, keyboard, descriptive files, location, sources, modes, signatures, etc.
- registries: ex. in mode class: script, charset, glyphs, pronunciation, etc.
- items: language, geographic location, architexter (architext management system), etc.
- etc.

They can be chained much like domain names, pointing a unique langtag per relational space, interlinking with other ISO 11179 conformant registries. They enable contexts to be built by inheritance from other referent or contextual registries.

- The target is set to identify 20,000 lingual spaces by the end of 2006. It will be associated with open use registration tools, ontology management programmes, registry/resolver logic, and proposed registry files.

- Each CRC is the local IANA of a language relational space. This is an indication about the size and complexity of a Multilingual Internet, whether using the MDRS approach or not.

INTF and MLTF

The MDRS structure, network solutions, and language roots are key building blocks for a Multilingual Internet by way of open usage. They are community projects that aim at:

- transparently and consistently extending the scope, surety, and proximity of the referent functions.
- enabling and supporting transparent cross-technology practices and usages.
- distributing linguistic communities empowerment, fighting the language divide, and balkanisation.

To build and maintain them, Intlnet serves as the secretariat of two specialised Task Forces under bootstrap:

- the International Network TF (**INTF**) for MDRS and network architecture, testing, sources, and support.
- the Multilingual TF (**MLTF**) for every language related issue.

Thank you

MDRS introduction	http://intlnet.org/mdrs.pdf
INTF	http://intf.org
MLTF	http://mltf.org