Proposal for an MoU/MG statement concerning

Semantic Interoperability

and the need for

a coherent policy for a framework of distributed, coordinated repositories for all kinds of content items on a world-wide scale

Standards in the context of eBusiness can mean among others
- state-of-the-art or standard practices in eBusiness
- industry standards
- international standards in the field of eBusiness.

The following statement provides an outline of the complexity behind the term semantic interoperability and then exemplifies the impact of semantic interoperability focused on cultural diversity on standards in all of the meanings above.

1

Under the perspective of **ubiquitous and pervasive computing**, technology should gradually disappear behind content, the presentation of content and user-friendliness. ‘Soft’ aspects, like culture and emotion, increasingly influence trust and need to be considered in information design. This will be reinforced in the course of development towards the convergence of **mobile computing and mobile communication** (MCC), when **eContent** develops in the direction of **mContent** etc. This development also necessitates **multi-channel** approaches without media-breaks in ICT. Furthermore, at political level the issue of **accessibility** (incl. the requirements of people with special needs) is gaining more attention at national (e.g. CLF* in Canada) and international (e.g. IFAP* of UNESCO) levels.

Tim Berners Lee’s conception of the **Semantic Web**, therefore, needs some extensions from the point of view of future mContent. Only then it can develop into the **global content infrastructure** for eBusiness, eLearning, eHealth, eGovernment, and other e…s, for which it had been conceived. In order to be efficient and effective, this generalized semantic web must provide **rules and procedures as well as organizational frameworks** to guarantee or at least support **different kinds of interoperability**, such as technical, operational and semantic interoperability:

- throughout the enterprise/organization,
- between enterprises/organizations,
- within industry consortia,
- between industry consortia (which urgently needs open standards),
- among different e…s,
- between different language communities,

which requires many new (incl. new types of) methodology standards.
First of all, certain fundamental methodology standards valid for all application fields (viz. the e...s) need to be developed, which will pave the way for semantic interoperability under the requirements of

- multilinguality
- cultural diversity
- multimodality
- accessibility (incl. the requirements of people with special needs)
- multi-channel presentations.

All of them comprise to a larger or lesser degree ‘soft’ aspects, which have to be considered at the earliest stage of software design before implementation. In this connection it must be clarified that semantic interoperability can/must be further subdivided into

- syntactic interoperability
- conceptual interoperability (incl. terminology, language resources, ontologies, etc.)
- pragmatic interoperability (comprising also the aspects of cultural diversity, etc.)

Beside, basic methodology standards on the basis of the above-mentioned fundamental methodology standards, which are specific to certain broad application fields, will ensure semantic interoperability within the same application area.

Furthermore, at the age of the semantic web (which has already evolved beyond the original conception of Tim Berners Lee), computers have to communicate among themselves in seemingly natural language, which – contrary to “true” natural language – has to be highly unambiguous. The developing information society, therefore, will need a large number of registries for many different types of repositories for

- certain types of data items/information objects: authority data, attributes, values, etc.,
- certain data elements, metadata/data categories, etc.
- terminological data of all sorts,
- typologies, taxonomies, nomenclatures, ontologies, etc.,
- coding systems (such as for names of countries, currencies, languages, airports, etc.)
- non-linguistic representations of knowledge (such as graphic symbols, etc.),
- data structures/datamodels & metamodels,
- interchange formats,
- XML schemas,
- syntactic communication protocols, messages, etc.,
- all kinds of conversion routines,
- all kinds of interfaces,
- other kinds of data dictionaries (containing data items/information objects),
- other kinds of data dictionaries (containing metadata, etc.),
- etc.

whose updating and maintenance has to be taken care of by Maintenance Agencies (MAs), Registration Authorities (RAs) and all kinds of Registries for these repositories.
This will require a systematic approach to the establishment of such MAs, RAs and Registries and calls for a policy of the world-wide standardization system, to establish a coherent framework for such registries, duly taking into account:

- different types of data items/information objects to be taken care of by the respective MAs, RAs and registries
- the degree of authoritativeness of each type of item/object
- the proper objectives of standardized and non-standardized updating/maintenance procedures
- the terms of reference of these MAs, RAs and registries
- different business models for different types of these registries (dealing with different kinds of items/objects)
- the work methodology as well as workflow management methods to be used in the updating/maintenance process
- legal aspects (comprising among others copyright and other IPRs)
- etc.

4

In this connection, terminology and other language resources (LRs) play a particular role, insofar as they are not only constituting elements of all kinds of ‘textual’ (incl. alphanumeric) content, but also serve as indexing and retrieval means for virtually all kinds of content. In addition, terminology serves as a means of knowledge presentation and knowledge ordering at the level of conceptual logic, while language resources also are indispensable at the user interface for converting spoken into written language and vice versa. If properly maintained terminology and other language resources can serve as a core tool for quality management of all kinds of content.

| A coherent policy for a distributed, however, well coordinated framework for all kinds of content items today only exists in a rudimentary form. The development may well end up in a network of distributed (federated) MAs, RAs and registries becoming the backbone of the e/mContent infrastructures of the semantic web. In this connection, the standards bodies not only will find new opportunities for standardization activities, but also have the societal responsibility to take the lead in organizing this coherent framework. |

**Abbreviations:**
- CLF – Common Look and Feel
- IFAP – Information for All Programme
- LR – language resource
- MA – maintenance agency
- RA – registration authority
- XML – eXtensible Markup Language