

Semantic Web: Global Unique ID for all Copyright Material

LATE Abran Kra Florence*, Shinji Kono†

*University of the Ryukyus/Departement of Information Engineering

†University of the Ryukyus

Abstract—The Semantic web will allow for significantly machine- readable content to be available on the world wide web. Getting this content onto the web, and using it once it is there, requires a global unique identifier for each data. In this paper, we introduce the implication of technology identifying each copyright material exclude machine generating one by a single globally unique identifier: GUID. This technology allows internet content providers straightforwardly to improve the current web search service.

I. INTRODUCTION

How often do you need access to electronic information, be it an article, an image, a bibliography or a video, that you do manage to find it, how do you know wether it is the current version and what the copyright implications are for using any of the content? In the fast changing world of content production, keeping track of the location and ownership of electronic files is difficult. Wouldn't it be really useful if you could link to the specific content you are searching for and also be notified of the copyright implications for its use with a single unique Id. That is the challenge facing researchers involved in the development of GUID.

A. Problems

The idea of a Uniform Resource Locator (URL), or hyperlink, is well known from the World Wide Web . The URL essentially says where on the web a target resource is located. The WWW mechanism of specifying a file by its URL is inadequate for ensuring the consistency and currency of mirrored copies, as a URL for an independently mirrored copy of a software package may point to an out-of-date copy and give no indication that it is not up-to-date. Futher-more, mirror copies of a file cannot be located from a URL reference, since each copy has a different URL. Maintaining the quality of software and of indexing information and presenting a uniform searching become much more difficult. Consistency between a set of files that are meant to be used together must be maintained. The main problems of URL itself are:

- The location of the ressources; the files get moved, changing their URLs.
- URL is dangling; files can leave a forwarding address at the old URL.
- URL is not unique; you have to scan the documentations to verify that compatible pieces have not been retrieved.

Most of the above problems can be solve by assigning a unique global Id for each copy right material when registered for worldwide use.

In this paper, we describe our research efforts of GUID for each copy right material in which we describe the GUID implementation

II. IMPLEMENTATION REQUIREMENT

To issue unique identifiers efficiently on a global scale, using only one registration authority to centralize and manage all identifiers is not reliable. We need a layer structure of issuing agencies distributed around the world just like the internet domain name system. There is one worldwide registration authority, which authorizes the original agency in each region, like a country, by assigning a Region Code to each agency. Region agencies authorize a GUID issue centres by assigning ID Center Number to them. Each ID issue center assigns a Intra-center Number to each copy right item to identify it. The combination of the region code, ID Center Number and Intra-center Number and defining the naming authorities responsible for naming each element, the identifier can keep its uniqueness as well as being issued efficiently. In the layer of distributed identifier issuing system shown in figure1, each element of identifier is issued by a different naming authority, which means each element can indicate its origin. This enables efficient distributed resolution service provision.

Consideration of the number of to-be-issued shows that one ID issue center for one region(country) is not enough. If there are multiple ID center in one region, one should take the lead role. The minimum, and optical, layered system consists of world registration authority, regional agency and ID issue center.

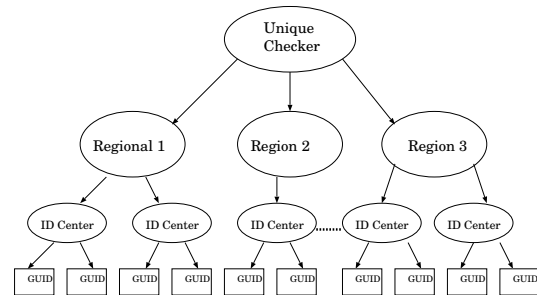


Fig. 1. Layer Confi guration of Identifi er issuing organisation.

Region code: is the number assigned by the registration Authority to each regional agency.

ID Center number: is the number assigned by each regional agency to ID issue center who issue Identifiers.

Intra-center number: is the identification assigned by each ID center to each digital item. the same digital item with exactly the same bits may be the subject of multiple actions, transactions and Users. In such a case , different numbers may be assigned to each combination of digital item and Action/Transaction/Users. The issuing policy of this part of the identifier depends on each ID issue center.

III. LINKAGE BETWEEN GUID AND OTHERS PHYSICAL OR SOCIAL ORGANISATIONS

This section describe the relation between the GUID registration/assignment and others physical organisations as shown in figure2.

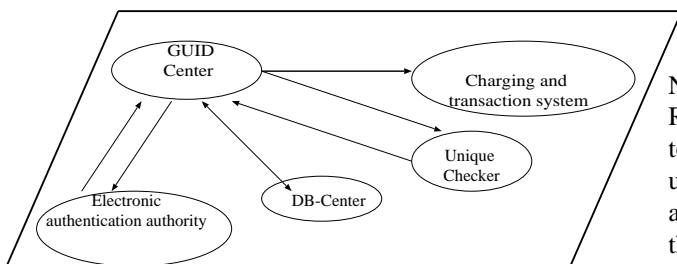


Fig. 2. Physical Organisations.

GUID management center: An organisation which manages the functions of both a GUID issuing center and an DB-center.

DB-Center: An organisation which stores and manages in a database attribute information uniquely identifying and describing items or packages of digital content which have been granted a GUID by a GUID issuing center.

Registration Authority: An organisation, which issues unique identifiers to content ID issuing centers, determines the rules under which content ID issuing center operate and ensures that these rules are adhered to.

IV. ARCHITECTURE OF SEMANTIC WEB INCLUDING GUID SERVER

Metadata or information about data is very important when it comes to handling unstructured data as text. The public sector has a very special need for standardising the descriptions of the huge amount of textual information that today resides in a multitude of data bases and archives in various, decentralised and geographically dispersed institutions. A global unique identifier assigned to each instance of content metadata set will describe a content and its distribution attributes.

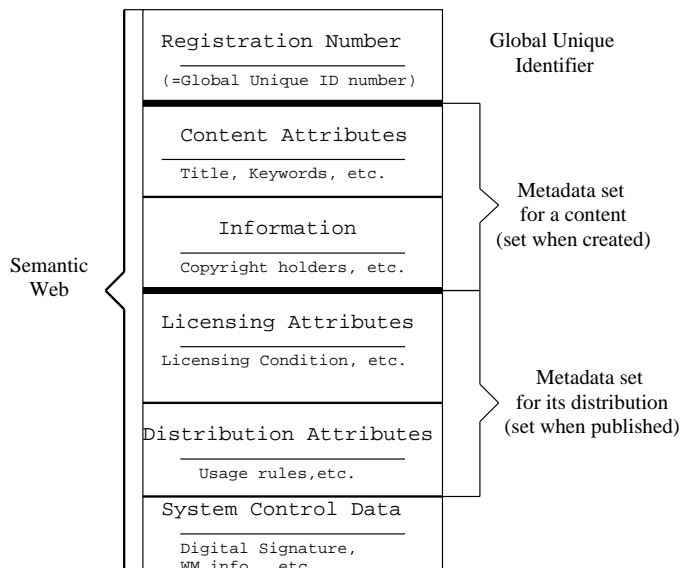


Fig. 3. Architecture of SW including GUID Server.

V. RELATED WORK

Functional requirements for Uniform Resource Names(URNs) are propose in [5] by IETF Uniform Resource Identification(URI) working group. According to [5], the function of a URN is to provide a globally unique, persistent identifier used for recognition of and for access to characteristics of a resource or to the ressource itself. URN assignment is delegated to naming authorities, the names of which are persistent and globally unique, and who may assign names directly or delegate their authority to sub-authorities. Global uniqueness of URNs is guaranteed by requiring each naming authority to determine the conditions under which it will issue a URN (for example, whether or not to issue a new URN when the contents of a file change).

VI. CONCLUSION

The Semantic Web requires new tools that can be used in new ways. One important use will be the GUID, allowing people to dynamically create and use Semantic Web information. Building such system will solve the URL's problems, and we describe some ideas aimed at providing this basis. Thus, the registration method described in this paper are examples of some of the basic technologies that are needed to create the Internet Next Generation.

REFERENCES

- [1] R. R Larson. Design and development of a network-based electronic library. in *Proc. ASIS Mid-year Meeting, pages 95-114, Portland, Oregon, May 1994.*
- [2] S. Browne and J. Dongarra. Location-Independent Naming for virtual Distributed Software Repositories.
- [3] H. Sakamoto, H. Fuji and J. Kishigami Digital Item Identification and Description based on Content ID Technology.
- [4] J. W.Moore. The use of encryption to ensure the integrity of reusable software components in *Proc. Third International conference on Software Reusability. IEEE Computer Society Press, Nov. 1992.*
- [5] K. Sollins and L. Masinter. Function requirements for Uniform Ressource Names. in *Internet RFC 1737, Dec. 1994*