



Brussels, 11 June 2015
(OR. en)

9191/15

LIMITE

JURINFO 13
EJUSTICE 76

NOTE

From:	Publications Office of the European Union
To:	Working Party on e-Law (e-Law)
Subject:	Metadata standardisation - State of play

1. Harmonised metadata is a very important factor when it comes to efficient and effective exchange of documents and information between different public services. Structured and harmonised metadata facilitate and improve the access to as well as the reuse and the exploitation of public information of the European Union.
2. The European institutions have already established synergies and coordinate their efforts in order to avoid duplication of work and to increase interoperability.
3. As a reminder, these efforts are coordinated by the Interinstitutional Metadata Maintenance Committee (IMMC), in charge of common metadata definition, encoding rules and exchange protocols to be used by the European Union institutions and bodies. The IMMC gathers as members all the institutions and is chaired by the Publications Office.
4. Several workflows based on the IMMC protocols have been implemented notably with regard to the transmission of documents to be published in the Official Journal. Bilateral discussions between the Publications Office and the institutions continue with a view to designing and introducing other channels of exchange, also based on the IMMC protocols.

I. Interinstitutional Formats Committee (IFC)

5. In the same spirit and in order to facilitate the production, exchange, publication, preservation and reuse of EU legislative documents, an Interinstitutional Formats Committee (IFC) has been created.
6. The role of the Interinstitutional Formats Committee (IFC) is to propose common semantic structures, in principle based on XML, for Institutions' documents, in particular legislative documents. The aim of the group is to formulate specifications that help improving interoperability of document flows between institutions, including validation rules for the common structure.
7. The IFC also serves as an interinstitutional platform for collaboration and knowledge exchange and guarantees the maintenance of the specifications. It further aims at supporting rationalization by exploiting the benefits of the gradual introduction of efficient methodologies.
8. A close collaboration between IMMC and IFC is in place, aiming at safeguarding consistency of the two standards.
9. In parallel, the Publications Office continues its work in the following areas of metadata standardisation:

II. Public procurement standardisation

10. The Publication Office works actively and directly with the Commission DGs GROWTH and DIGIT in international committees addressing standardization of public procurement. The clear ambition is to ensure, at least at the European level, that a standard is developed for public procurement data which is to be published and/or exchanged.
11. In parallel the Publications Office is developing data models describing the notices published on TED. These models will allow the streamlining of the standard forms used by the contracting authorities when communicating to the informants their procurement procedures that are to be published on TED. Data models are also being produced to cover the different systems and processes in the eProcurement domain to enhance interoperability between internal and external services.

12. The Publications Office is in charge of eTendering which is a system that helps contracting authorities launch their procurement procedures. The system is already used and will soon be used by an important number of Commission services as well as the European Parliament, the Court of Auditors, the European Investment bank and different European agencies.

III. "Court Data Base" project

13. In the framework of the Court Data Base project, led by DG Justice of the Commission, the Publications Office created a prototype of an authority table to include the Members States' national courts. This authority table (1) will be used to populate both the Court Data Base and sector 8 of EUR-Lex (national case law).
14. The ultimate goal of this project is to incorporate the Court Data Base to the European e-Justice Portal.

IV. ISA (Interoperability Solutions for Public Administrations)

15. The Publications Office is actively engaged in the Action 1.1 of the ISA activities which aims to promote the semantic interoperability amongst the European Union Member States.
16. An overview of the on-going project is given below:
 - a) *Permanent URIs and common approach for the management of URIs by EU institutions*
17. End of 2014 the URI task force delivered a "Study on a common policy for the governance and management of persistent URIs by EU institutions".
18. The study underlines the importance of a common URI policy among the EU institutions and public services in general and suggests a common approach on three topics: design rules (URI pattern), governance and management of the URIs.

¹ List of authorised metadata values used in the exchange of documents between the institutions.

b) Pilot project "Dissemination of the EU Budget as Linked Open Data"

19. This project started in April 2015 and its main purposes are to improve transparency of the EU public data and to facilitate access and interoperability. A first step includes the creation of a dissemination vocabulary which will allow the reuse of EU budget data.
20. For information, the EU budget is currently published in different formats such as HTML, PDF, XML (FORMEX) and will also be available shortly in Excel format on the EU Open Data Portal. This already allows for some processing and analysis, but not for the rich linking enabled by Linked Open Data.

c) Revision of the DCAT-AP application profile

21. As a reminder, the **DCAT Application profile for data portals in Europe** (DCAT-AP) is a specification based on the Data Catalogue vocabulary (DCAT) for describing public sector datasets in Europe. Its basic use case is to enable cross-data portal search for data sets and make public sector data better searchable across borders and sectors. This can be achieved by the exchange of descriptions of datasets among data portals ⁽¹⁾.
22. The DCAT-AP specification has been chosen as the metadata standard for the description of the datasets of the pan-European Open Data portal.

V. Management of thesauri and other controlled vocabularies

23. A corporate platform set up by the Publications Office will enable the management and dissemination of multilingual thesauri and other controlled vocabularies (e.g. the authority tables published in the Metadata Registry ⁽²⁾).
24. This corporate service will allow for: a) better indexation, b) easier access and retrieval of information and c) collaborative management of different thesauri.
25. Further elaboration on this issue will be brought under agenda item 11, EuroVoc.

¹ https://joinup.ec.europa.eu/asset/dcat_application_profile/description

² <http://publications.europa.eu/mdr/>