Evolving a NASA Digital Object Identifiers System with Community Engagement

Lalit Wanchoo$^1$ (lalit.wanchoo@nasa.gov) and Nathan James$^2$ (nathan.j.james@nasa.gov).

$^1$ ADNET Systems, Inc., 7515 Mission Drive, Suite A100, Lanham, MD 20706; $^2$ Earth Science Data and Information System Project (Code 423), NASA Goddard Space Flight Center, Greenbelt, MD 20771

Introduction

In 2010, NASA’s Earth Science Data and Information System (ESDIS) Project began investigating the assignment of unique identifiers to its suite of data products being stewarded at different data centers across the country. This led to the use of Digital Object Identifiers (DOIs) and the development of an automated system for the registration of DOIs.

A key factor in the successful evolution of the DOI registration system has been the incorporation of community input. Over the last three years, ESDIS solicited community input for making the DOI registration process more efficient from three focus groups under NASA’s Earth Science Data System Working Group (ESDSWG). These groups were largely composed of DOI submitters and data curators from the 12 data centers serving user communities of various science disciplines. The suggestions from these groups were formulated into recommendations for ESDIS consideration and implementation.

The ESDIS DOI registration system has evolved to be fully functional with over 3,500 publicly accessible DOIs and over 200 DOIs being held in reserve status until the information required for registration is obtained. The goal is to assign DOIs to the entire 7000+ data collections under ESDIS management via its network of discipline-oriented data centers.

Objective

To demonstrate how the ESDIS DOI system and its processes have evolved over these years based on the recommendations provided by the user community (whether the community members create and manage DOI information or use DOIs in the data citations) the user community is comprised of people with common interests and needs for data identifiers who are actively involved in the creation and usage process. Engagement describes the interactive context wherein the community provides information, evaluates the proposed processes, and provides guidance in the area of identifiers.

Benefits of Community Engagement

Development and implementation of processes, services, and systems that best serve the user community.

Community Engagement Approach

Selection (Identifier Scheme)

System Development

ESDIS established a system that addressed the needs of the community and provided a process that allowed identifiers to be reserved and registered. The operational system is fairly automated and uses a Perl script with an Oracle database for the data. The system functionality provides reports using Oracle Apex.

System Evaluation

The ESDIS System and its registration process has been primarily developed based on the input from the ESDIS community. This successful system, which has been in operation for the last four years, has processes over 5000 DOI requests from 11 data providers (see Figure 2). Currently, over 4,000 DOIs have been registered.

Acknowledgments

Lalit Wanchoo’s contributions to this study were funded through the Science and Exploration Data Analysis (SESDA III) GSFC NASA Contract No. NNG12PL17C and Nathan James’s contributions to this study were made as a part of his employment by NASA.

Figure 1: Components of ESDIS DOI Registration and Management System

Figure 2: Monthly DOI Requests Received by ESDIS for Registration

Table 1: Suitability Assessment of Selected Identifier Schemes

| Identifier Scheme | Full Name | URI | DOI | RFID | PMID | PURL | URL | UUID | File | Handel | Handle System | SOI | CGI | ARK | LSID | GSID | UID | Summary
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PURL</td>
<td>Persistent Unique Resource Identifiers</td>
<td>Persistent Unique Resource Identifiers</td>
<td>Persistent Unique Resource Identifiers</td>
<td>Persistent Unique Resource Identifiers</td>
<td>Persistent Unique Resource Identifiers</td>
<td>Persistent Unique Resource Identifiers</td>
<td>Persistent Unique Resource Identifiers</td>
<td>Persistent Unique Resource Identifiers</td>
<td>Persistent Unique Resource Identifiers</td>
<td>Persistent Unique Resource Identifiers</td>
<td>Persistent Unique Resource Identifiers</td>
<td>Persistent Unique Resource Identifiers</td>
<td>Persistent Unique Resource Identifiers</td>
<td>Persistent Unique Resource Identifiers</td>
<td>Persistent Unique Resource Identifiers</td>
<td>Persistent Unique Resource Identifiers</td>
<td>Persistent Unique Resource Identifiers</td>
<td>Persistent Unique Resource Identifiers</td>
</tr>
<tr>
<td>GSID</td>
<td>Geospatial Identifiers</td>
<td>Geospatial Identifiers</td>
<td>Geospatial Identifiers</td>
<td>Geospatial Identifiers</td>
<td>Geospatial Identifiers</td>
<td>Geospatial Identifiers</td>
<td>Geospatial Identifiers</td>
<td>Geospatial Identifiers</td>
<td>Geospatial Identifiers</td>
<td>Geospatial Identifiers</td>
<td>Geospatial Identifiers</td>
<td>Geospatial Identifiers</td>
<td>Geospatial Identifiers</td>
<td>Geospatial Identifiers</td>
<td>Geospatial Identifiers</td>
<td>Geospatial Identifiers</td>
<td>Geospatial Identifiers</td>
<td></td>
</tr>
<tr>
<td>UID</td>
<td>Universally Unique Identifiers</td>
<td>Universally Unique Identifiers</td>
<td>Universally Unique Identifiers</td>
<td>Universally Unique Identifiers</td>
<td>Universally Unique Identifiers</td>
<td>Universally Unique Identifiers</td>
<td>Universally Unique Identifiers</td>
<td>Universally Unique Identifiers</td>
<td>Universally Unique Identifiers</td>
<td>Universally Unique Identifiers</td>
<td>Universally Unique Identifiers</td>
<td>Universally Unique Identifiers</td>
<td>Universally Unique Identifiers</td>
<td>Universally Unique Identifiers</td>
<td>Universally Unique Identifiers</td>
<td>Universally Unique Identifiers</td>
<td>Universally Unique Identifiers</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Current Practices Recommendations

Table 3: Landing Page Recommendations

Table 4: Data Citation Recommendations

Usage and Feedback

To promote consistency, discoverability, and usefulness across NASA, the ESDISW established three working groups to analyze the implementation and usage of DOIs. These groups recommended the landing page and data citation requirements for the data products that are distributed by ESDIS. Such recommendations, listed in Tables 3 and 4, were accepted and implemented. The Data Citation group used the Mandatory and Suggested content for citations given in the Data Citation Guidelines for Data Providers and Archives document developed and approved by the Federation of Earth Science Information Partners (ESIP) Federation. These guidelines shall be followed to the greatest extent possible in developing citations for ESDIS data products.

Conclusion

The ESDIS DOI System and its registration process has been primarily developed based on the input from the ESDIS community. This successful system, which has been in operation for the last four years, has processed over 5000 DOI requests from 11 data providers (see Figure 2).

In addition, the group evaluated the information being made available to the people who are assigning DOIs to the products. Recommendations from this group and the action taken by ESDIS are shown in Table 2. These recommendations were accepted and implemented.