

# URNs in the Wild

Ron Daniel, Jr.

Taxonomy Strategies LLC

[rdaniel@taxonomystrategies.com](mailto:rdaniel@taxonomystrategies.com)

# Agenda

- Introduction
- Definitions
- Identification
- Namespace Creation
- Namespace Registration
- Resolution Mechanisms
- Resolver Deployment
- Conclusions

# Introduction

- Non-goals:
  - Advocating for or against URNs or any other particular identification scheme.
  - Advocating for or against any specific resolution mechanism.
- Goal of this talk:
  - Describe easily-visible aspects of current URN usage in the world.



From <http://www.global-insight.com/Enlightened.htm>

# Definitions

- URN – A string obeying the syntax:

**urn:***namespaceID:namespace-specific-string*

- *namespaceID* may be registered or unregistered

# Agenda

- Introduction
- Definitions
- Identification
- Namespace Creation
- Namespace Registration
- Resolution Mechanisms
- Resolver Deployment
- Conclusions

# Identification

- How many resources are identified by publicly-resolvable URNs?
  - No good estimation method found - Impression: “Not large”
- How many resources **use** URNs to identify other resources?
  - Estimate: Large and growing
  - urn:schemas-microsoft-com:office:\*
    - Google reports 53M public MS Word docs containing “the” indexed in the last 3 months.
    - Adding private documents, other languages, longer timeframe, other MS formats ->
    - **On order of 1B Office documents using URNs today to identify something**
  - Other uses as well – Open Office, xmlns=“urn:\*”, several others
- Most uses are “boring” – no dynamic retrieval
  - Does that matter?

# Namespace Creation

## ■ Registered Namespace – one of:

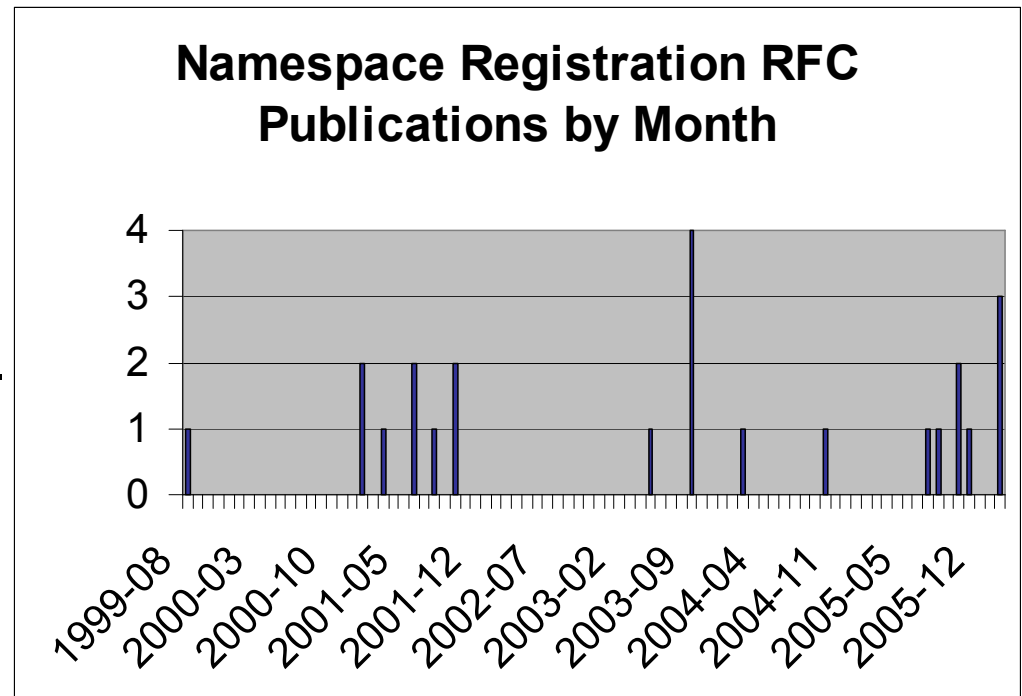
IETF	OASIS	NBN	Fipa	UUID	Fdc
PIN	XMLORG	WEB3D	Swift	UCI	ISAN
ISSN	Publicid	MPEG	Liberty	CLEI	NZL
OID	ISBN	Mace	IPTC	Tva	Oma
NEWSML					

- Unregistered namespace – any other *namespaceID*
  - **Unregistered namespaces dominate the observable URNs**
  - Some interesting unregistered schemes (e.g. EPC\*).
  - Many small examples.

\* EPC is “in the process of being registered” – M. Mealling

# Namespace Registration

- A relatively heavyweight, but free, registration process is in place.
  - It is being used, but not frequently (25 registered namespaces).
  - Unregistered namespaces **dominate** observable URNs.



# Resolution Mechanisms

- **No universally-applicable URN resolution software exists.**
  - Several namespaces were explicitly intended to NOT be globally resolvable.
- Resolution is application-dependent and/or namespace-dependent.
- The **most commonly-used resolution mechanisms**, in order, appear to be:
  1. None
    - e.g. MS Office
  2. XML Catalogs
    - e.g. OASIS, MPEG
  3. DDDS/NAPTR
    - e.g. EPC's ONS resolution service, ISBN (informally)
  4. Scheme-specific dynamic resolution
    - e.g. PIN
- NAPTR resolution also used for things other than URNs
  - e.g. ENUM - but VOIP puts ENUM's growth in doubt.

Namespace	RFC	Resolver
IETF	2648	Custom
PIN	3043	Custom
ISSN	3044	Underspecified
OID	3061	Unspecified
NEWSML	3085	Delegated to assigners
OASIS	3121	<b>Catalogs</b>
XMLORG	3120	<b>Catalogs</b>
publicid	3151	Various, including Catalogs
ISBN	3187	Underspecified
NBN	3188	Underspecified
WEB3D	3541	<b>Catalogs</b>
MPEG	3614	<b>Catalogs</b>
mace	3613	Website list
fipa	3616	Website list
swift	3615	Custom
liberty	3622	Unspecified
IPTC	3937	TBD
UUID	4122	N/A
CLEI	4152	Custom
UCI	4179	Custom
tva	4195	N/A
fdc	4198	Delegated
ISAN	4246	TBD
NZL	4350	Website list
oma	4358	N/A

# Resolver Deployment

## ■ XML Catalogs

- Broad support in XML software.

## ■ NAPTR

- Support is currently built into BIND and some other DNS servers.
- 3'rd party extensions exist for browsers, but not widely used.

# Agenda

- Introduction
- Definitions
- Identification
- Namespace Creation
- Namespace Registration
- Resolution Mechanisms
- Resolver Deployment
- **Conclusions**

# Conclusions

- Unsurprisingly, there is little dynamic resolution of URNs to Resources.
  - Primary use is identifying ‘known’ info, specifically XML Namespaces.
- Surprisingly, many more resources *use* URNs than are identified by URNs *themselves*.
- Most uses are “boring” – no dynamic retrieval
  - Does that matter?
- ***URNs are being used differently than URLs!***
  - Pure identification, even w/o resolution, has value.