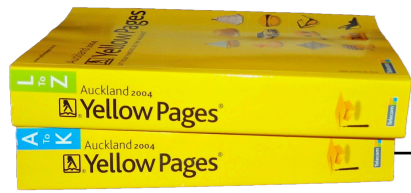


Report on Data Providers Onboarding Activities

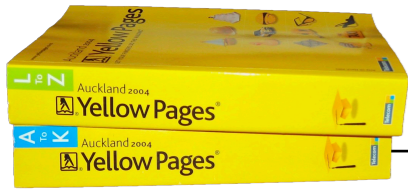
Renaud Savalle, Hendrik Heintz

2023-11-11 - Registry WG + Edu IG



Outline

- ▶ A bit of History
- ▶ The IVOA Registries
- ▶ The Registry Ecosystem
- ▶ Key points from “onboarding” online meeting
- ▶ Documentation and Tools for setting up a Publishing Registry



A bit of History... (2004)

- <https://ivoa.net/documents/Notes/RMExp/RMExperience-20040305.pdf>
- Lessons learned in using the **VOResource** XML schemas, focusing on what made using the schemas *difficult*...
- Conclusion: ...Can adoption of standard schemas be accelerated if they were revised to make them *easier to use* ?



*International
Virtual
Observatory
Alliance*

Lessons Learned Using the VOResource XML Schemas in the NVO Version 1.0

IVOA Note 5 March 2004

This version:

<http://www.ivoa.net/Documents/Notes/RMExp/RMExp-20040305.html>

Latest version:

<http://www.ivoa.net/Documents/latest/RMExp.html>

Previous versions:

none

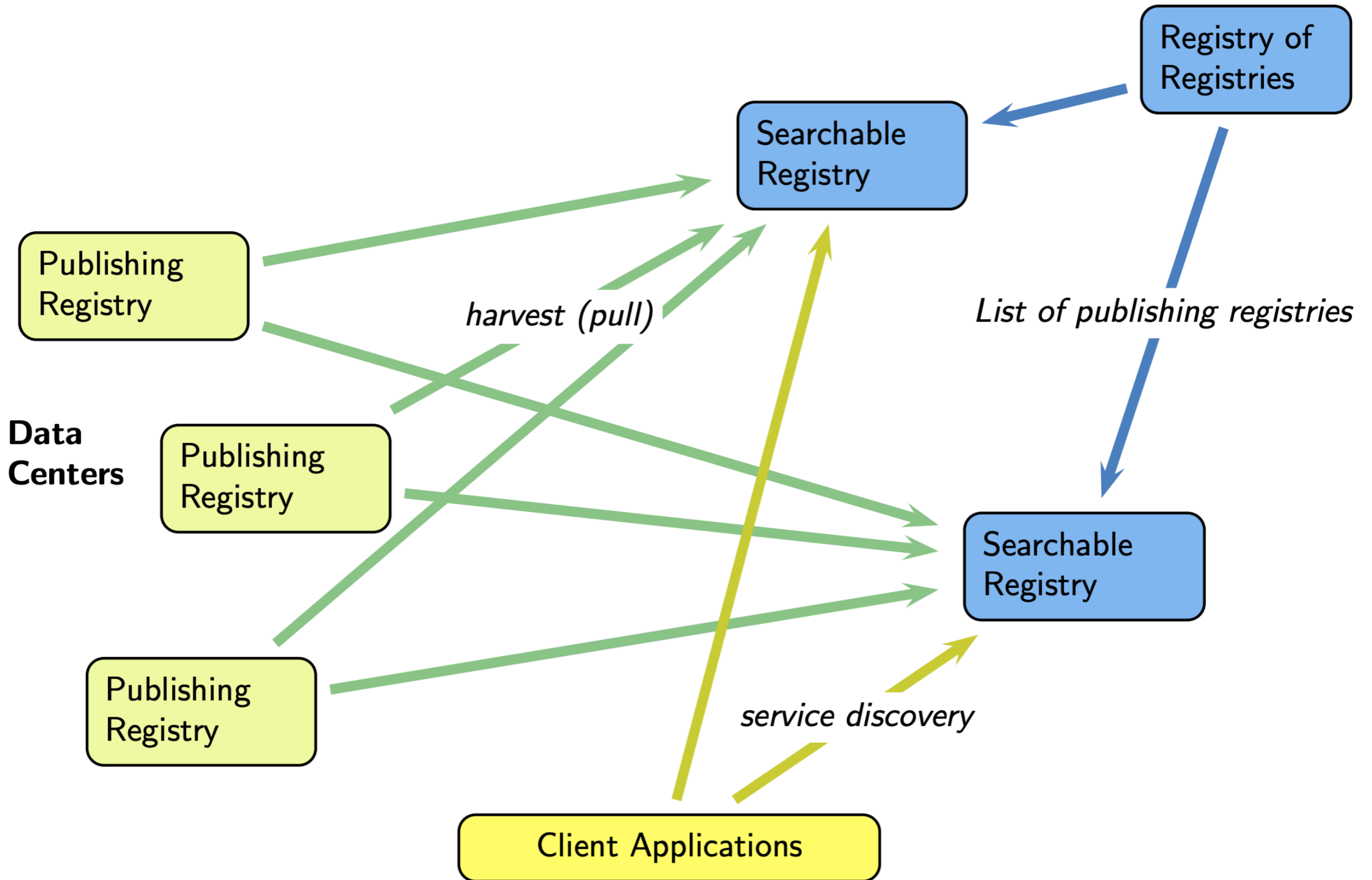
Authors:

[Raymond Plante](#) (NCSA)
[Matthew Graham](#) (Caltech)
[Gretchen Greene](#) (STScI)
[Bob Hanisch](#) (STScI)
Jeongin Lee (HEASARC)
[Wil O'Mullane](#) (JHU)
[Ramon Williamson](#) (NCSA)

Abstract

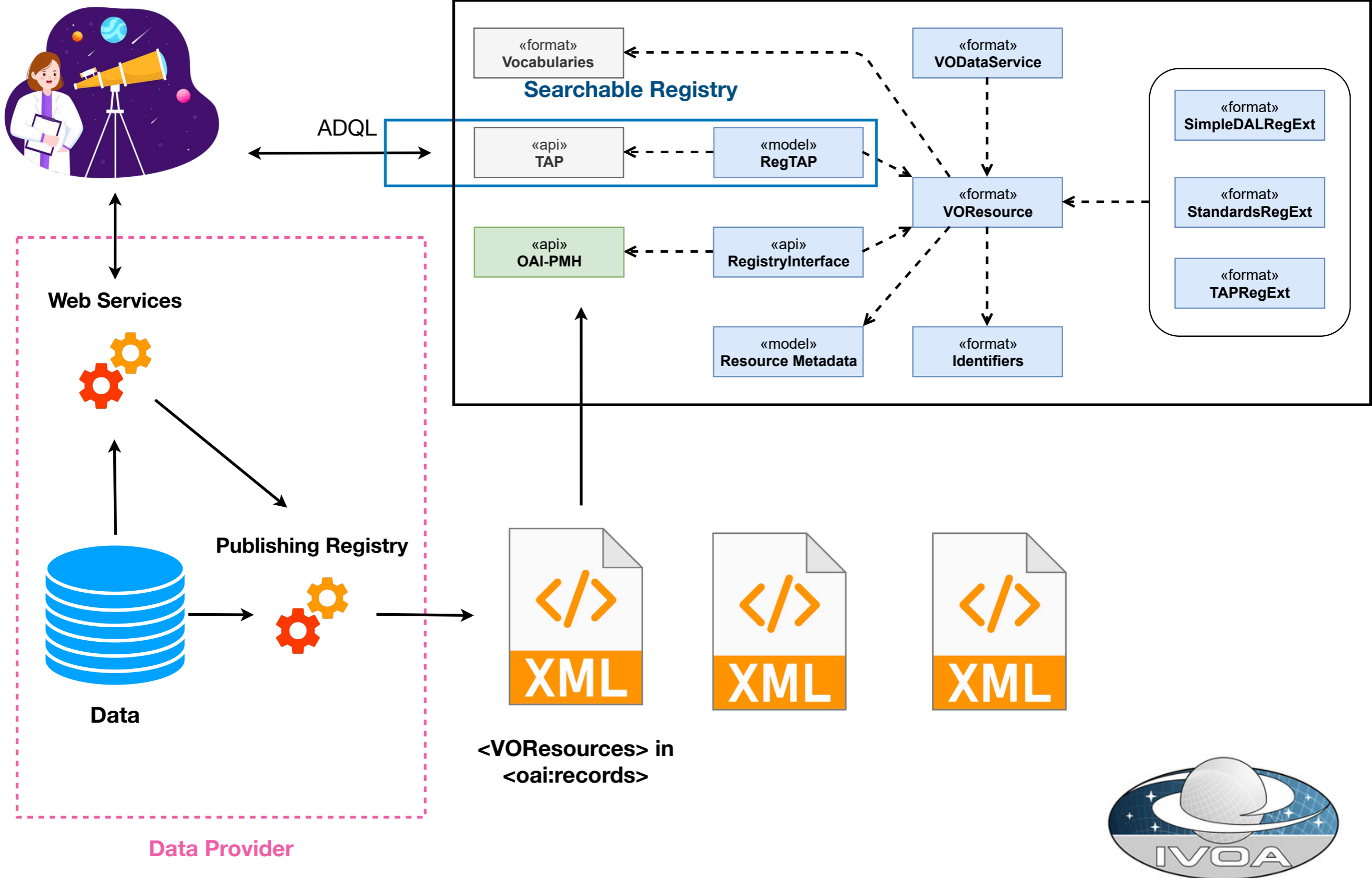
During 2003, the [National Virtual Observatory \(NVO\)](#) project in the US set out to prototype the VO Registry framework as described by the [IVOA Registry Working Group](#). A key goal was to demonstrate use of the [VOResource XML schemas](#) as a format for exchanging descriptions of resources in registries. While we feel that the effort was overall a successful validation of the framework and the XML schemas, it was particularly useful for understanding where the challenges lie. This document looks at our lessons learned in using the [VOResource XML schemas](#), focusing on what made using

The IVOA Registries

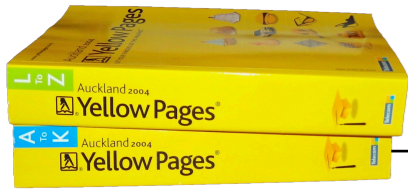


The Registry Ecosystem

Standards

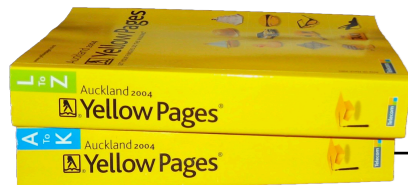


Registry of Registries



Key points from data providers “onboarding” online meeting (2023-10-03)

- Several established data providers shared their experience maintaining IVOA Publishing Registries. Thanks to participants from CADC, CDS/Vizier, CEFCA, GAVO, StSci/MAST !
- Using OAI-PMH interface in short: Publishing Registry maintainers need to
 1. Write/generate VOResources for their dataset and services
 2. Encapsulate them into oai:records and
 3. Provide a persistent OAI endpoint with micro-services:
Ex: https://cds.unistra.fr/registry/?verb=ListRecords&metadataPrefix=ivo_vor
(This is then used by Full Searchable Registries to incrementally harvest Publishing Registries)
- General advice: do not implement the OAI-PMH interface from scratch, since existing packages/libraries can be used in several programming languages. It is also possible to use frameworks (such as GAVO's Dachs) which can take care of publication in addition to service implementation.
- Validation of publishing registries is possible using the online tool used as the first step (Validating the Harvesting Interface) when adding a publishing registry on the RofR page (<http://rofr.ivoa.net/>). It was also advised to do XSL validation first.
- Meeting minutes with more details: https://docs.google.com/document/d/1_4SLsTQj6nbwcl1oqPyiCM4Rg8p1QFhA32_DWh6Rcl0 (or ask me the PDF)



Documentation and Tools for setting up a Publishing Registry

- Documentation
 - **Overview:** Paper “The virtual observatory registry” by Demleitner et al. 2014A&C.....7..101D
 - <https://wiki.ivoa.net/twiki/bin/view/IVOA/GettingIntoTheRegistry>
 - **Standards** ! (maybe start with VOResource and VODataService ?)
- Tools
 - Euro-VO Registry’s web "**wizard**" is adapted to publishers with a few number of manually published services.
 - PUblishing Registry proXy (**purx**) (<https://dc.g-vo.org/purx/q/enroll/info>) does not require to generate OAI-PMH (but you don’t get your own authority identifier)
 - **Libraries** implementing the OAI-PMH protocol (ex in Python: <https://pypi.org/project/pyoai/>)
 - Publishing Registry **Validator** <http://rofr.ivoa.net/regvalidate/>

An non-exhaustive list !

Euro-VO Registry Interface

VOResource creation web “wizard” + XML editor

EUROPEAN SPACE AGENCY ABOUT ESAC

Euro-VO Registry

MY RESOURCES

Results (45)

Information

Create minimal Table Access 1.0 Resource
(you may later use the forms interface to add or edit elements...)

Status*: active
Ivoid*: ivo://esavo/
Title*:
Publisher*:
Contact Name*:
Subject*:
Description*:
Reference URL*:
Language Name*:
Language Version Name*:
Output Format Mime*:
Access URL*:
Use*: full

OK Cancel

Short Name	IVOID	Status	Capabilities	Compliance	#total versions	Created
	ivo://esavo/gaia/conesearch	is dele			3	2018-09-29
	ivo://esavo/hst--ssap	is dele			2	2016-02-29
	ivo://esavo/isda/siap	is dele			1	2016-02-05
	ivo://esavo/iso/siap	is dele			10	2020-11-28
	ivo://esavo/iso/ssap	is dele			11	2020-11-28
	ivo://esavo/nist/slap	is dele			12	2019-09-20
	ivo://esavo/test3	activ			1	2022-12-06
	ivo://esavo/xsa/cs/XMMSL2	is dele			2	2018-10-30
3XMM-DR7-stacked	ivo://esavo/xsa/cs/3xmmdr7	activ			2	2019-05-16
3XMM-DR8	ivo://esavo/xsa/cs/3xmm	activ			9	2019-05-16
EHST/HLA/SIAP	ivo://esavo/hla/siap	activ			6	2019-05-16
EHST/HLA/SSA	ivo://esavo/hla/ssap	inact			4	2019-05-16
EHST/HST/SIAP	ivo://esavo/hst/siap	activ			5	2020-01-30
EHST/HST/SSAP	ivo://esavo/hst/ssap	activ			7	2020-01-30
ESA	ivo://esavo/esavo	activ			4	2019-05-16
ESA-ESDC	ivo://esavo	activ			4	2019-05-16
ESASky	ivo://esavo/esasky/tap	active	ESASky TAP vs:CatalogService	Partially Compliant(C--), not applicable, not applicable, not applicable	9	2023-01-09

Page size: 100

Displaying 1-45 of 45

<https://registry.euro-vo.org/evor/#info>

Onboarding the registry

An education matter

Hendrik Heint, Renaud Savalle, Markus Demleitner

November 10, 2023



Outline

Edu matters

Conclusion



Newcomers experience

As we were working on this in the past years, we receive similar feedback as before, and these are points we (as in IVOA) should address:

1. No single entry point to start from.
2. Standards documents are hard to understand for newcomers
3. Registry records are hard to validate
4. Information on the web pages are outdated



Hackathon Workshops for Newcomers

Following the Vespa example of 1 week workshops for newcomers with the goal to have a prototype service with **real** (their) data running by the end of the workshop.



‘European Data Provider Forum’

During the European collaboration projects (Cosadie, Asterix, ESCAPE) we had meetings at which we invited data providers to talk about their experiences and solutions, but also to train them on existing software solutions to publish their data. We think a concept similar to this would be a good outreach opportunity.



Test the services

Curating the registry records also is something we should raise awareness to. It's crucial that services are not only up and "running" but also are well maintained and curated. "Buggy" services mess with the User experience hence we do have an interest to identify them, and either support their repair, or shut down.



House calls

Visit existing data providers to support them with problems on their services. Use the opportunity for also doing a VO-workshop.



Establish an entry point

Yes, that's a rather empty phrase, I know. Yet it's worth thinking about what we can set up and maintain: maybe a webpage ? A subcommunity on github ? Discord Channels ? Regular meetings between Interops, for counseling, Q&As and Complaints management ?



Lot's to do.

All these efforts sum up to quite a bit of workload, which will not be easy to commit to, especially since most are on positions dedicated to different work. Henc we are looking for....



Funding

Money, money, money
must be funny
in the rich man's world

– Swedish proverb

