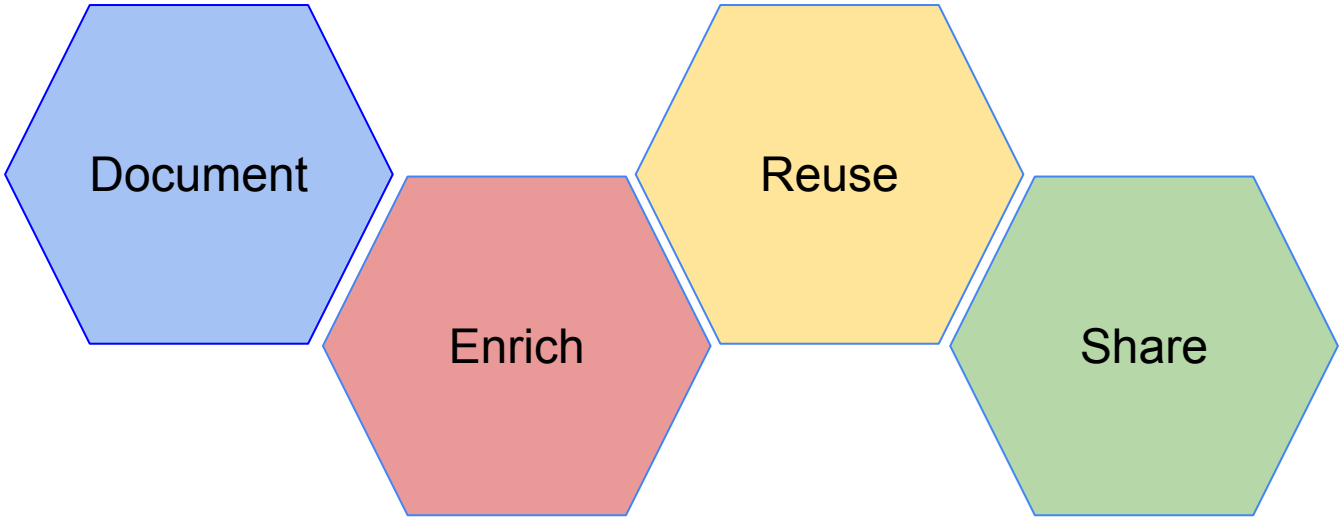
Zellij - the semantic pattern library

supporting sustainable semantic data management

George Bruseker (Takin.solutions)
CIDOC CRM SIG
26/09/2024

State of the Art

Zellij: Semantic Data Pattern Management



Document

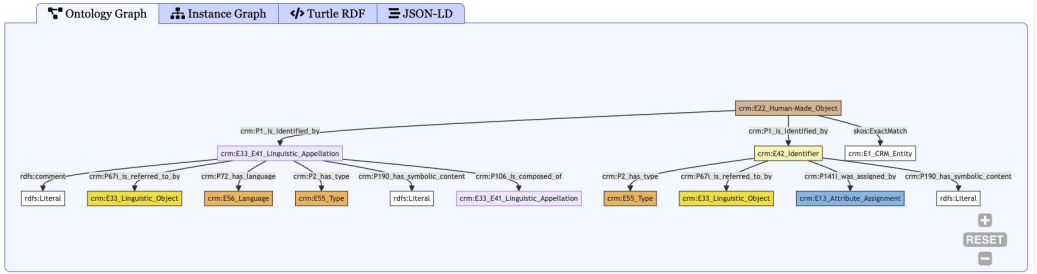
Reusable, Understandable Semantic Patterns

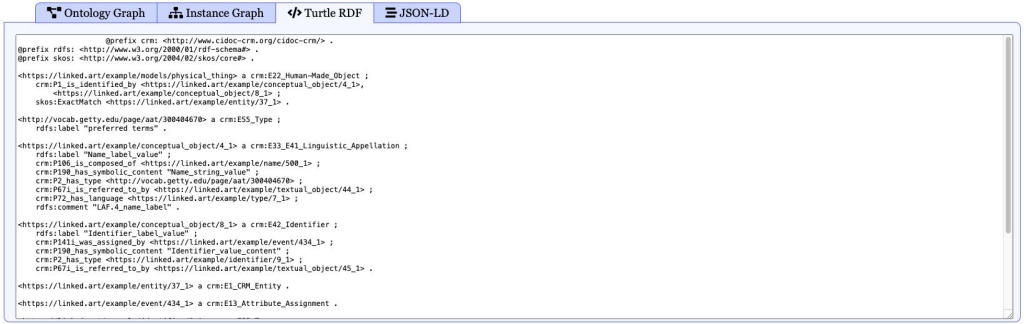
Models
Collections
Fields

Documentation

[CAT.1] Names and Identifiers

Identifier	Name	Field System Name	Description	CRM Path	Expected Value Type	Expected Collection	Expected Model
LAF.6	Name	name_content	This field is used to record the string value of the name attributed to the documented physical thing.	->p1->E33_E41[4_1]->p190->rdf:literal	String		
LAF.5	Name Type	name_type	This field is used to record the type of the name attributed to the documented physical thing.	->p1->E33_41[4_1]->p2->E55[5_1]	Concept		
LAF.7	Name Language	name_language	This field is used to record the language of the name attributed to the documented physical thing.	->p1->E33_E41[4_1]->p72->E56[7_1]	Concept		
LAF.44	Source Reference Work for Name	name_source_reference	This field is used to link to a source text in which the name denoting the documented physical thing is used.	->p1->E33_E41[4_1]->p67i->E33[44_1]	Reference Model		Textual Work
LAF.4	Name Label	name_label	This field is used to record the string value of the machine readable label used for displaying the instance of name that is used to denote the documented physical thing.	->p1->E33_41[4_1]->rdfs:label->rdf:literal	String		
LAF.500	Name Part	name_part	This field is used to link the documented name of an entity to its relevant name part.	->p1->E33_41[4_1]->p106->E33_E41[500_1]	Collection	Name	





Standard Mapping Patterns

Enrich

Manage
Control
Sustain
Reuse

ETL Processes

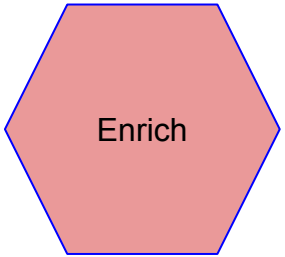
Fields

Identifier	Name	Description	Field System Name	CRM Path	Expected Value Type	Mapping Pattern
LAF.6	Name	This field is used to record the string value of the name attributed to the documented entity.	name_content	->p1->E33_E41[4_1]->p190->rdf:literal	String	<link template="LAF.6_name_content"> <path> <source_relation> <relation/> </source_relation> <target_relation> <relationship>crm:P1_is_identified_by</relationship> <entity variable="4_1"> <type>crm:E33_E41_Linguistic_Appellation</type> <instance_generator name="UUID"/> </entity> <relationship>crm:P190_has_symbolic_content</relationship> </target_relation> </path> <range> <source_node/> <target_node> <entity> <type>http://www.w3.org/2001/XMLSchema#string</type> <instance_generator name="Literal"> <arg name="text" type="xpath">text()</arg> <arg name="language" type="constant">en</arg> </instance_generator> </entity> </target_node> </range> </link>
LAF.5	Name Type	This field is used to record the type of the name attributed to the documented entity.	name_type	->p1->E33_41[4_1]->p2->E55[5_1]	Concept	<link template="LAF.5_name_type"> <path> <source_relation> <relation/> </source_relation> <target_relation> <relationship>crm:P1_is_identified_by</relationship> <entity variable="4_1"> <type>crm:E33_E41_Linguistic_Appellation</type> <instance_generator name="UUID"/> </entity> <relationship>crm:P2_has_type</relationship> </target_relation> </path> <range> <source_node/> <target_node> <entity variable="5_1"> <type>crm:E55_Type</type> <instance_generator name="UUID"/> </entity> </target_node> </range> </link>
LAF.7	Name Language	This field is used to record the language of the name attributed to the documented entity.	name_language	->p1->E33_E41[4_1]->p72->E56[7_1]	Concept	<link template="LAF.7_name_language"> <path> <source_relation> <relation/> </source_relation> <target_relation> <relationship>crm:P1_is_identified_by</relationship> <entity variable="4_1"> <type>crm:E33_E41_Linguistic_Appellation</type> <instance_generator name="UUID"/> </entity> <relationship>crm:P72_has_language</relationship> </target_relation> </path> <range> <source_node/> <target_node> <entity variable="7_1"> <type>crm:E56_Language</type> <instance_generator name="UUID"/> </entity> </target_node> </range> </link>
LAF.44	Source Reference Work for Name	This field is used to link to a source text in which the name denoting the documented entity is used.	name_source_reference	->p1->E33_E41[4_1]->p67i->E33[44_1]	Reference Model	<link template="LAF.44_name_source_reference"> <path> <source_relation> <relation/> </source_relation> <target_relation> <relationship>crm:P67i_is_referred_to_by</relationship> </target_relation> </path> <range> <source_node/> <target_node> <entity variable="44_1"> <type>crm:E33_Linguistic_Object</type> <instance_generator name="UUID"/> </entity> </target_node> </range> </link>

Click to edit domain

Click on a row to edit the matching table

#		SOURCE	TARGET PATH NAME	TARGET	IF RULE	COMMENTS
1	D		LAP.1_name	E33_E41_Linguistic_Appellation		
1.1	P		LAF.5_name_type	*P2_has_type*		
	R			E55_Type [5_1]		
1.2	P		LAF.6_name_content	*P190_has_symbolic_content*		
	R			XMLSchema#string		
1.3	P		LAF.7_name_language	*P72_has_language*		
	R			E56_Language [7_1]		
1.4	P		LAF.44_name_source_reference	*P67i_is_referred_to_by*		
	R			E33_Linguistic_Object [44_1]		
1.5	P		LAF.500_name_part	*P106_is_composed_of*		
	R			E33_E41_Linguistic_Appellation [500_1]		



Standard Query Patterns

Retrieve
Count
Visualize
QA

Query Functions

Fields

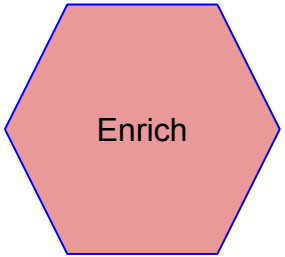
Identifier	Name	Description	Field System Name	CRM Path	Expected Value Type	Query Pattern
LAF.6	Name	This field is used to record the string value of the name attributed to the documented entity.	name_content	->p1->E33_E41[4_1]->p190->rdf:literal	String	?subject crm:P1_is_identified_by ?LAP_1_name . ?LAP_1_name a crm:E33_E41_Linguistic_Appellation . ?LAP_1_name crm:P190_has_symbolic_content ?LAF_6_name_content . BIND(?LAF_6_name_content as ?value) .
LAF.5	Name Type	This field is used to record the type of the name attributed to the documented entity.	name_type	->p1->E33_41[4_1]->p2->E55[5_1]	Concept	?subject crm:P1_is_identified_by ?LAP_1_name . ?LAP_1_name a crm:E33_E41_Linguistic_Appellation . ?LAP_1_name crm:P2_has_type ?LAF_5_name_type . BIND(?LAF_5_name_type as ?value) . OPTIONAL { ?LAF_5_name_type rdfs:label ?LAF_5_name_type_label . }
LAF.7	Name Language	This field is used to record the language of the name attributed to the documented entity.	name_language	->p1->E33_E41[4_1]->p72->E56[7_1]	Concept	?subject crm:P1_is_identified_by ?LAP_1_name . ?LAP_1_name a crm:E33_E41_Linguistic_Appellation . ?LAP_1_name crm:P72_has_language ?LAF_7_name_language . BIND(?LAF_7_name_language as ?value) . OPTIONAL { ?LAF_7_name_language rdfs:label ?LAF_7_name_language_label . }
LAF.44	Source Reference Work for Name	This field is used to link to a source text in which the name denoting the documented entity is used.	name_source_reference	->p1->E33_E41[4_1]->p67i->E33[44_1]	Reference Model	?subject crm:P1_is_identified_by ?LAP_1_name . ?LAP_1_name a crm:E33_E41_Linguistic_Appellation . ?LAP_1_name crm:P67i_is_referred_to_by ?LAF_44_name_source_reference . OPTIONAL {?LAF_44_name_source_reference rdfs:label ?LAF_44_name_source_reference_label . }
LAF.4	Name Label	This field is used to record the string value of the machine readable label used for displaying the instance of name that is used to denote the documented entity.	name_label	->p1->E33_41[4_1]->rdfs:label->rdf:literal	String	?subject crm:P1_is_identified_by ?LAP_1_name . ?LAP_1_name a crm:E33_E41_Linguistic_Appellation . ?LAP_1_name rdfs:label ?LAF_4_name_label . BIND(?LAF_4_name_label as ?value) .
LAF.500	Name Part	This field is used to link the documented name of an entity to its relevant name part.	name_part	->p1->E33_41[4_1]->p106->E33_E41[500_1]	Collection	?subject crm:P1_is_identified_by ?LAP_1_name . ?LAP_1_name a crm:E33_E41_Linguistic_Appellation . ?LAP_1_name crm:P106_is_composed_of ?LAF_500_name_part . BIND(?LAF_500_name_part as ?value) . OPTIONAL {?LAF_500_name_part rdfs:label ?LAF_500_name_part_label . }

SPARQL Endpoint /census_HI/ Content Type (SELECT) JSON Content Type (GRAPH) Turtle

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX crm: <http://www.cidoc-crm.org/cidoc-crm/>
SELECT * WHERE {
 ?subject crm:P1_is_identified_by ?LAP_1_name . ?LAP_1_name a crm:E33_E41_Linguistic_Appellation . ?LAP_1_name crm:P190_has_symbolic_content ?LAF_6_name_content . BIND(?LAF_6_name_content as ?value) .
} LIMIT 10
```

	subject	LAP_1_name	LAF_6_name_content
1	<https://semantic.census.de/physicalthing/43956/event/production/timespan>	<https://semantic.census.de/physicalthing/43956/event/production/timespan/appellation/125D9CBB-326B-...	post 1538:1:1-ante 1571
2	<https://semantic.census.de/physicalthing/43956/dimension/53A1324A-5E5D-4F23-...	<https://semantic.census.de/physical_thing/43956/dimension/appellation/73D08AFF-3FDA-3175-BED9-2A...	390 mm X 270 mm ca.
3	<https://semantic.census.de/physicalthing/61146>	<https://semantic.census.de/physicalthing/61146/appellation/2F6FB18C-C5CA-3795-8E48-AA74B845ACD4>	inv. 28-1-20
4	<https://semantic.census.de/physicalthing/61146>	<https://semantic.census.de/physicalthing/61146/appellation/B62EED99-FEED-3CBD-9D2B-ACAF05EB82D...	Codex: inv. 28-1-20

Standard Implementation Patterns



Arches
ResearchSpace
Metaphactory

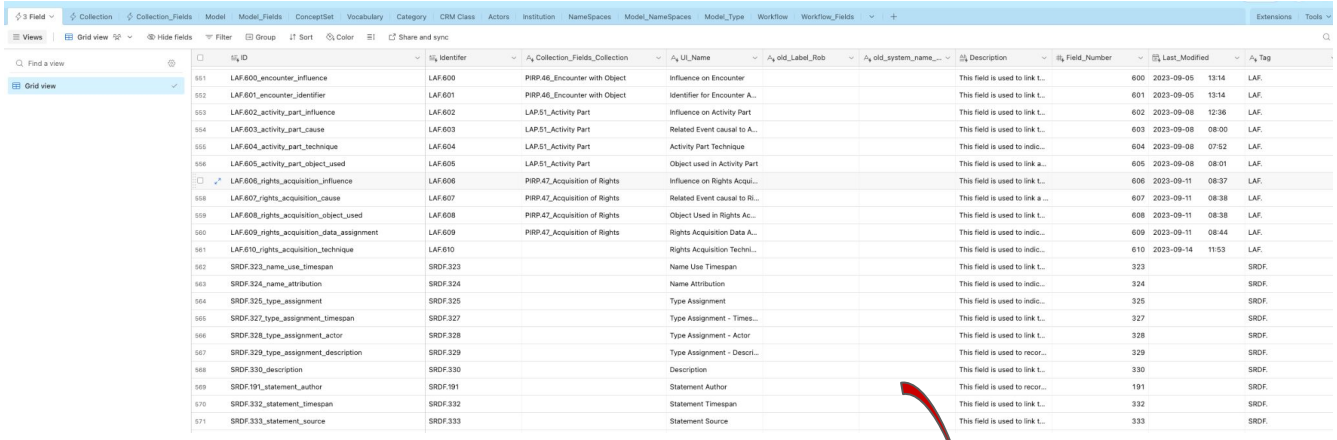
Ready Made Design

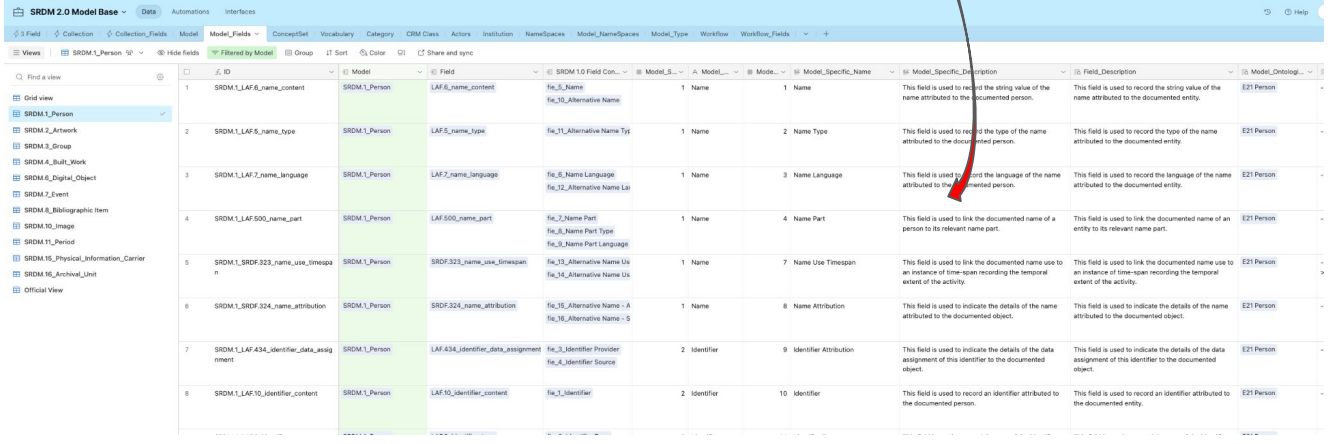
Reuse

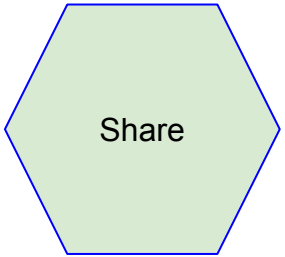
Reuse

Modularity
Expandability

Reduce-Reuse-Recycle







Share

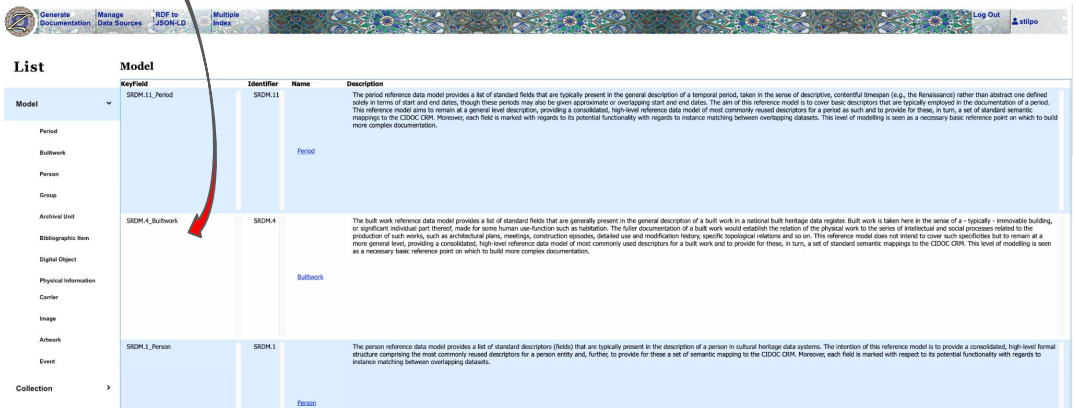
Partnership
Interoperability
Efficiency

A Semantic Pattern Library

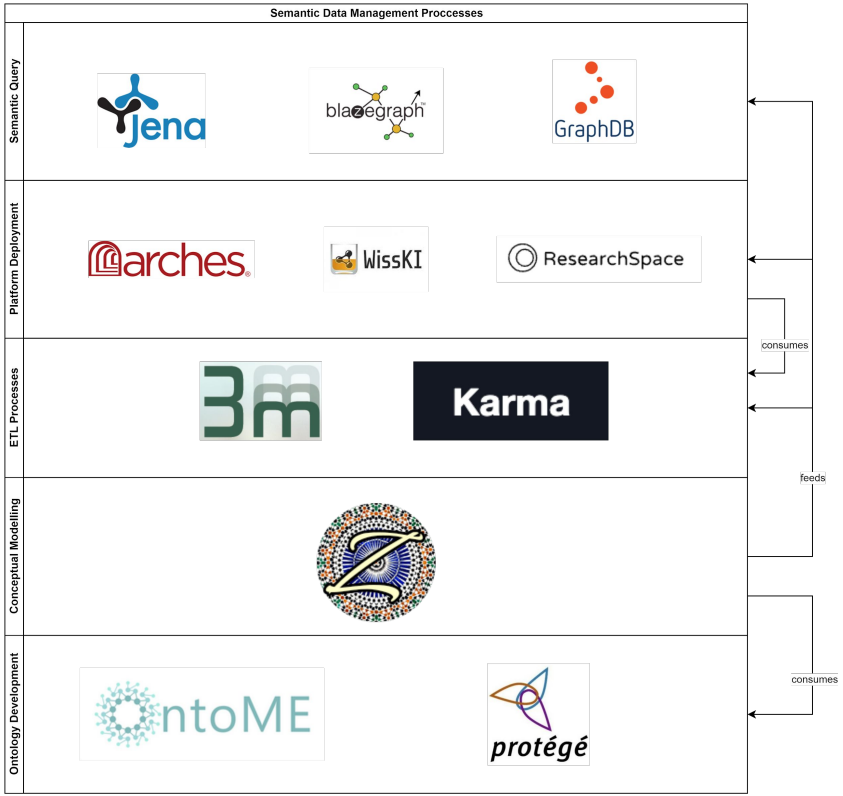
Pick an ontological database

Search: Search... Search

- SARI/SRDM 1.0
- Takin/Parks
- Takin/LW ITA
- Takin/Census
- Takin/Semafora
- Getty Semantic/Rusha Semantic
- Takin/SIA
- Getty Semantic/Provenance Index
- Takin/Zellij Demo
- SARI/JILA
- Takin/DHI
- Takin/PDLM
- Getty Semantic/Linked.Art
- SARI/SKKG
- SARI/SRDM 2.0



Zellij in the Semantic Stack



Why? What's the standard workflow?

Zellij based projects don't start from scratch

All previous modelling work from previous projects is reusable

Build out a semantic data model strategy faster and have the documentation ready **from the start**

Your project builds out of the documentation so it always follows the specification

Mapping files are pulled from the patterns

QA devices like SParQL and SHaCL are pulled from the patterns

Templates to launch your new wisski, researchspace, arches platform are also from the patterns

One source to rule them all and in the darkness bind them. Ie no more infinite loops of what did the design do, what did we actually implement, what did the ETL team do, how does it compare with the plan?

The Community so far

SARI [SRDM Project]

Parks Canada {National Database of Recognized CH]

LifeWatch [Datasets Database]

Humboldt University [Census Project]

University of Groenigen [CRM Survey]

Getty Digital [Provenance Index, Arches for Science, Linked.Art]

German Institute in Rome [Slavery Contracts Database]

Max Planck History of Science [Projects Database]

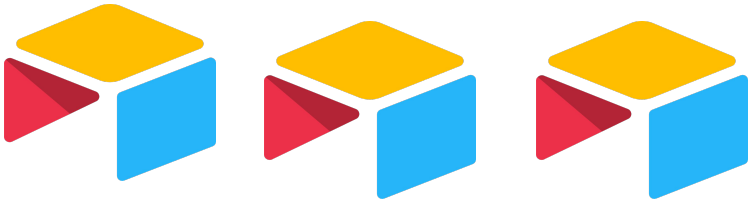
Getty Conservation [Lingo Vocabaulary Project]

Philadelphia Museum of Art [Art Information Commons]

Present Architecture

Public UI
Delivered by Jinja
Python templates

Python based Jinja templates custom code reads Airtable API



Airtable as Database
And Editor UI

Moving Forward - Spiffing it up!

Spiffy - Semantic Pattern Interchange Format (for you)

1. XML Exporter creates system neutral xml files describing
    a. Project
    b. Composite Pattern
    c. Atomic Pattern
2. Spiffy Patterns Versioned and publicly available in Github
3. Spiffy files used to automatically generate derivatives from patterns:
    a. RDF
    b. SparQL
    c. YML Setup File for ResearchSpace
    d. X3ML file for 3M
    e. Arches JSON LD model file for Arches

Looking for Synergies…

Spiffy - Semantic Pattern Interchange Format (for you)

1. CRM SIG
    a. Open to
        i. being an official place to document recommended patterns
        ii. providing a data standard for sharing patterns
    b. Could benefit from
        i. Collaboration to identify needs for xml interchange format and development there around
2. Ontological Mapping Software
    a. Would love to call on an API to pull the latest version of an ontology
    b. Would love to call on an API to check what is out of data in an existing project/space (set of paths)
3. Semantic Mapping Software
    a. Can already map to RDF/XML, need to be able to create JSON-LD in context in order to serve Arches
    b. Would like to create ShaCL generators to check validity of patterns
4. Semantic Platform Software
    a. Open to all systems to building a generator for interface generation (e.g. Arches models and branches, Wisski branches)
5. Educational Software
    a. Would like to further integrate CRITERIA to document fields in diagrams
    b. Would like to integrate with tools like OMG and drawio to triples to support teaching / testing scenarios e.g. via ShaCL patterns to indicate a good an bad response