Iriomote cat illustration by Prof. Mitsuru Moriguchi, Okinawa University.







Why we do what we do?

- Institution with 11 facilities
- over 40 Million exhibition and research items
- Very heterogeneous data and data infrastructures:
- Common database: https://search.senckenberg.de currently 1.521.698 objects in 124 collections
- Facility and department maintained implementations 2.
- 3. Locally stored data

Conclusion: Only a partial FAIRness is given with a potentially wide range of existing and upcoming items and collections.

BITS Terminology Service Overview

- Terminologies, specifically for RDF and OWL ٠
- Access for humans and machines
- Curation of new terms (vocabulary) .
- Metadata management
- Visualization of semantic artifacts



Search for ontology, te		roperties and individualis	Search		
-	EXALMAL N	country lens		H	
Filter		Browse Ontologies			
		page size 10 w	sorted by Title		
By heyword		ABCD Base Ontology	64 Cannon		
		The base cetology of the ABCD standard	404 Properties		
ollection		Collections: FAR Data Spaces, 532	Looded. 2023-02-02		
(c) internet	501	Destabless filefor	2145 Oatorto		
O OwtoPLANT	0	Germon decimal classification system based on the Dulch Basic Class	stratos (IRC) 22 Procedes		
S ADD	0	Collections: FID BAUdgild, FID more, MPD4NG, ESS			
D RPOIACHEM			TEL MARKER ALL	-	
O PID mile	0	Bioinformatics operations, data types, formats, identifi-			
D NFDIAGAT	0	EDAte is a single catology of well emploited, tambar corcep- bioinformatics, including types of data and data identifiers, data for		š	
D FAR Data Spaces	e e	controlmance, including types or easis and easis elements, easis to ECALI provides a set of terms with symphotics and definitions - organise for comparison use.		ð	
FID BAUEges	ő	Contestioner und Charlos Inforecterine 633	2022년 22년	-	
O CIERN	¢	CONTRACT OF CONTRACT.		1	
		Service Port		Т	



Preparation and integration

- Process and identify relevant data fields 1.
- Identify relevant elements such as words and phrases 2.
 - safely remove irrelevant content such as numbers a.
 - remove redundancy to improve performance b
- Use BITS TS to obtain metadata about the elements 3.
- Process unresolvable phrases by splitting them into 4. words
- Selecting the best results 5.
- Store the resulting data set in an appropriate manner for 6. human and machine agents (in progress)

Results and lessons learned

- Using pre-FAIR data sets requires a lot of manual work
- Identifying "the best result" (of over 30.000) for each element is not possible within an appropriate time frame: Using coverage based approach for the whole data set

Conclusion

Based on GO FAIR, we are already following most of the principles using the SGN infrastructure and BITS TS. Full implementation of the BITS TS extension to the SGN repository will provide a long-term strategic approach to providing machine-actionable digital specimens and enhancing FAIRness.

SENCKENBERG Alexander Wolodkin

world of biodiversity

https://orcid.org/0000-0003-0351-6523 https://orcid.org/0000-0003-1556-8750

Claus Weiland

This work is licensed under CC BY 4.0 (c)