

LifeBlock for FAIR data citation

Joaquín López Lérida Data e-Science Management Plans and Blockchain Officer, LifeWatch ERIC

https://orcid.org/0000-0001-9697-7710

Mastering Data Citation: Insights from the BioDT Research Infrastructures Online, 20, Nov 2023



Funded by the European Union



We needed to be able to support the most common biology metadata formats. These must include EML and DataCite in order to cater for the needs of non-biological metadata if needed and in the case a provider doesn't support EML.

Metadata Format	GBIF	LTER	ZENODO	METADATA CATALOG
EML	Х	Х		Х
DataCite	Х		Х	Х

A statistical analysis of the prevalence of the different fields across some RI files was done for metadata to have a clear idea about design needs.



Frequency of EML terms one level deep into the Metadata XML for some RIs.

Our challenges for LifeBlock metadata schema



- Rls Diversity: Each organization has its own language and structure; we had to translate and unifies these languages.
- **Multilingualism and Formats:** Multiple languages and formats, with linguistic and technical barriers.
- **Text to Knowledge:** Our goal is to convert unstructured information into valuable knowledge, ready to be explored and applied, particularly to our VREs.
- Interoperability: We need that VREs and RIs from different sources and formats can "talk" to each other seamlessly. Definition of a user data space where we can unify the different sources of information.
- Queries: Need of intuitive interface for querying and exploring RIs, both in federated syntactic search notation and in semantic and artificial intelligence search notation.





Metadata management for FAIR data citation in LifeWatch ERIC: LifeBlock

Federated and semantic Search over main RI's

Common metadata schema for RDF transformation

		Dataset Search		Meta	2 m						
					Betweet Course	mi, Gespaphi Temponi Senicos Wolflows Liktilok Prajet Pesple Acces apr Contage Countage Countage	ss ögnisalar lifustvicke fölgenet Bölggely Aclifes Amotalons lifetics. Other Testhilition Baix				
		LILEBIOCK	Federated search Semantic k	eyword search Semantic Structured Search	0	ladavet Title: Ceta	stigger of type spectrees of vescalar plants depended in the N		Installation ID:	2x335a56-7591-4/05-8ckc-a1bb77aa4ea6	
		Ψ				wheel Decrytion Inc.	, datest sectors the partmen recents from the statistice of L., , Checklet		Organization Key	6c4a0880-2a4d-1168-aa2d-88a03c50a862	
		User Tools	Import from		0	Meethpe 8	Analysis Analysis Family and Analysis		Type	IPT_INSTALLATION	
		My data	Select a data source	•		instandate 12.	27 11 / 2016		Title	GBIF Spain IPT	
		Search Tool	Select a data source		0	Mont IPt 80	Mitala		Description	This tool helps to GBIF Spain to do a tracking, increase the Data Qu	
		LifeBlock Explorer	GBIF				Advess Just		Modified By:	Tegery-myraton.got.org	
		Wallet Info	Zenodo			0	Adappa Mighe Magai		Created	20/02/2013	
		BESU Blockchain Network Manager	Lter		0	isator DR Had	sharing the Diministrated de Granada		Modified	19/07/2021	
		IPFS Network Manager	Metadata catalogue		0	onad Point UR:			Disabled:	~	
		Dev Tools	Re DiSSCo			hireLasta http	where gliopinsation to 403-406-900-4090		Contacts		
		Lifeblock Smart Contract Tool	N AI	h parameters		wren _			Endpoints	https://ipt.gbil.es/iss.do	
		Lifeblock API	Rediam	in parameters.		e 101	(S-riverse)		Machine Tags:		
			LifePortal		0	Ingenization Herd Namence	Baric de la Brisenidad de-Granata		Tags		
		Management loois	Eait selected	iport selected Import all		laguero: faintananos			Identifiers		
		Token Generator			0	teoption			Comments:		
		Emost Contract Monitor									
oken data	Traceability	Transfer token						-γ			
n 57			In the second seco								
Token id:	57		Heading Chew run he	wareny							
Block number:	119069				4		Usage of NFT (EF	(C-721) to	o sto	rage metadata	
Transaction has	ih: 0:819365 🖸			53				, , , , , , , , , , , , , , , , , , ,			
Timestamp:	2 months ago						Information as union	que and t	raza	ble information	
Contract addre	ss: <u>0xa1.9942</u> 🖸			52	•	C	ontainer with IPES a	s storage	resc	purce for datase	ste
Owner address	0xaf. 23b8 🗖					0		o otorugo	1000		,10
Title:	02_Catalogue of type	specimens of vascular plants deposited in the Herbarium of the Uni	iversity of Granada (Spain)	<u>63</u>			Each token	is a data	set t	hat has	
Dataset id:	708			tooot trooophility			an uniquo num	hor and a	n un	iquo ownor	
First ancestor:	53		Da	taset traceability			an unique num	Jei anu a	ii uii	ique owner.	
Type	NFT				_						
Token:	LWToken										
Symbol:	LWT										

Dataset control information

View all tokens

LifeBlock and FAIR data citation



- Findability Provide user-friendly search functions, intuitive for any user and above all, easy to use.
- Accessibility Provide accessible environments for users and easy to understand and use, even for semantic search.
- Interoperability Provide data beyond the resources generated by LifeWatch ERIC by extending the databases accessible from the tool to other RIs, DB, etc.
- **Reusability** Provide data from LifeBlock that can be used in any other environment, workflow, SKG, VRE, ..., always preserving their **traceability** in order to be guaranteed with successive reproductions.





- **Recognition and Attribution of Data**: LifeBlock's transparency and ERC-721's unique identification capabilities ensure precise tracking and attribution of each dataset, enhancing recognition of data creators and promoting open data sharing.
- Standard Practices in Data Citation: ERC-721 tokens enable clear tagging and citation of datasets, creating a standardized and respectful data citation practice, ensuring proper acknowledgment of creators and sources.
- **Improvement in Data Discovery and Reusability**: Lifeblock's immutable record, combined with ERC-721's detailed metadata, improves the findability and reusability of datasets, accelerating biodiversity research.
- **Credit to Data Managers**: LifeBlock highlights the contributions of data curators and managers, transparently demonstrating their vital role in maintaining and updating datasets.
- **Collaboration with Research Infrastructures**: Blockchain integration with research infrastructures like GBIF, Zenodo, LTER or Rediam and LifeWatch fosters seamless data sharing and enhances data citation practices through collaborative expertise.
- **Essential Techniques for Data Citation**: The usage of ERC-721 in data citation introduces key techniques for precise and efficient data referencing, offering a solid framework for recognizing contributions in biodiversity research.