

Where to with it?

Deciding where to publish data

Jan Vališ 

Dpt. of Analytical Chemistry, Faculty of Chemical Engineering, UCT Prague
& Academic Services, NTK



Co-funded by
the European Union

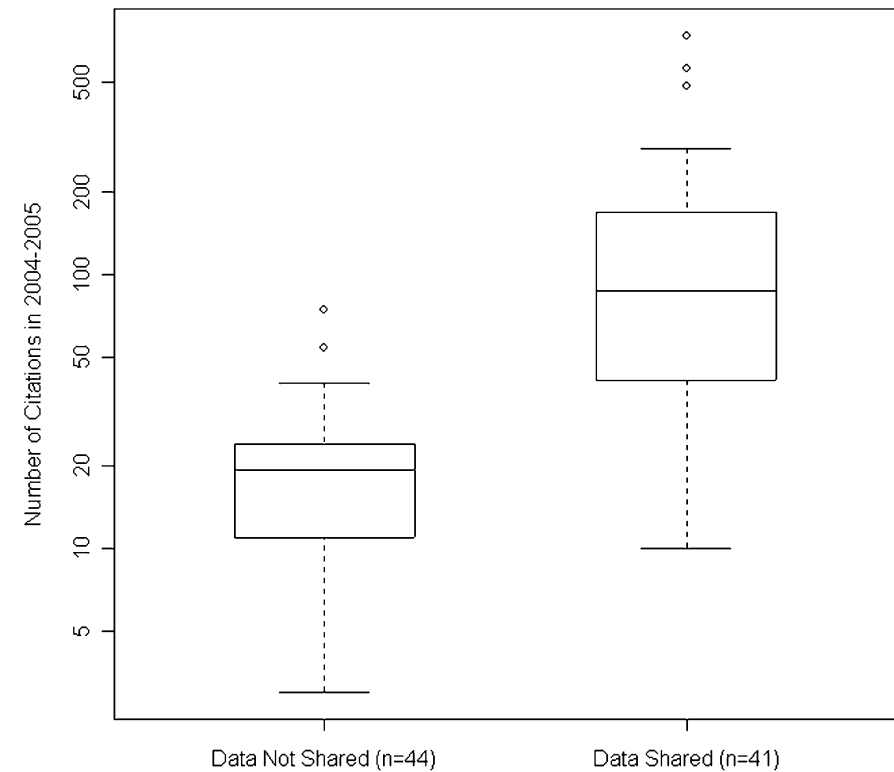


Who are you?

- Position (PhD candidate, postdoc, data-steward...)
- Field (Natural science, engineering ...)
- Do you have any experience with research data management?

Why should I care about a repository?

- Safety and security
- Clear licensing
- Transparency, integrity
- Sharing, collaboration
- Impact
- Compliance



2004–2005 citation counts of 85 trials by data availability.
Heather A. et al. 2007. PLOS ONE. License: [CC BY 4.0](#)
[10.1371/journal.pone.0000308](https://doi.org/10.1371/journal.pone.0000308).

Repositories

Repository types

- Institutional (e.g., CERN, ESO, ASEP)
- General purpose (e.g., Zenodo, National Repository)
- Field-specific/Disciplinary
 - Instrument/analysis/data type (e.g., MS, FT-IR, XRD)
 - Topic (e.g., drug discovery)

Repository types

- Position in lifecycle
 - raw primary data repository
 - processed data repository
 - **supporting data for an article**
- By curation
 - open
 - curated
 - peer-reviewed



What to ask when looking for a repository?

- Motivation? (community, publication or grant compliance?)
- Metadata & ontologies?
- File format compatibility?
- File size/number limitation?
- Access control/embargo?
- Licenses?
- Reputation, persistence, sustainability?
- Price?

Ideal characteristics

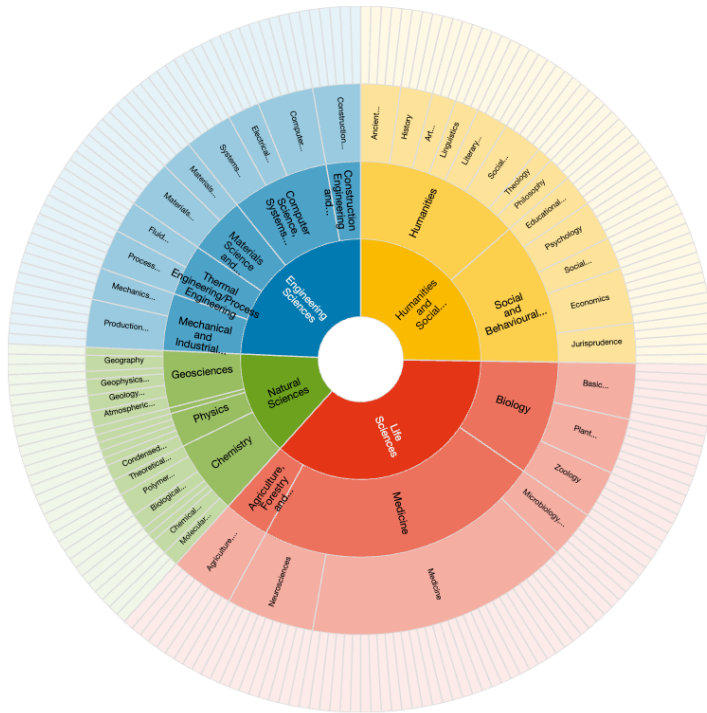
- Relevant metadata and ontologies
- PIDs (DOI, ORCiD, ROR, IGSN ...)
- Supports versioning
- FAIR
- Grant dedications
- Stable funding
- Free to contribute to

Red flags and pitfalls

- HTTP, expired certificates
- No clear sustainable business plan/funding
- Low number of records
- Publisher-owned
- No grant dedication
- No PIDs
- Irrelevant licenses

Where to look for repositories?

- All-purpose aggregators (e.g. [re3data](#), [FAIRsharing](#))
- Societies and grant agencies (e.g. [NFDI4Chem](#), [ERC](#))
- Publishers?



Illinois Data Bank

Humanities and Social Sciences Life Sciences Natural Sciences Engineering Sciences

institutional

dataProvider

United States

1. *Study*

University

[Kommunikation und Organisationsentwicklung](#)
[Kommunikationspsychologie](#)
[Kommunikationswissenschaft](#)
[Organisationspsychologie](#)

INSTRUCTIONS

4.8-10-11-12-13-14-15-16-17-18-19-20-21-22-23-24-25-26-27-28-29-30-31-32-33-34-35-36-37-38-39-40-41-42-43-44-45-46-47-48-49-50-51-52-53-54-55-56-57-58-59-60-61-62-63-64-65-66-67-68-69-70-71-72-73-74-75-76-77-78-79-80-81-82-83-84-85-86-87-88-89-90-91-92-93-94-95-96-97-98-99-100-101-102-103-104-105-106-107-108-109-110-111-112-113-114-115-116-117-118-119-120-121-122-123-124-125-126-127-128-129-130-131-132-133-134-135-136-137-138-139-140-141-142-143-144-145-146-147-148-149-150-151-152-153-154-155-156-157-158-159-160-161-162-163-164-165-166-167-168-169-170-171-172-173-174-175-176-177-178-179-180-181-182-183-184-185-186-187-188-189-190-191-192-193-194-195-196-197-198-199-200-201-202-203-204-205-206-207-208-209-210-211-212-213-214-215-216-217-218-219-220-221-222-223-224-225-226-227-228-229-230-231-232-233-234-235-236-237-238-239-240-241-242-243-244-245-246-247-248-249-250-251-252-253-254-255-256-257-258-259-260-261-262-263-264-265-266-267-268-269-270-271-272-273-274-275-276-277-278-279-280-281-282-283-284-285-286-287-288-289-290-291-292-293-294-295-296-297-298-299-300-301-302-303-304-305-306-307-308-309-310-311-312-313-314-315-316-317-318-319-320-321-322-323-324-325-326-327-328-329-330-331-332-333-334-335-336-337-338-339-340-341-342-343-344-345-346-347-348-349-350-351-352-353-354-355-356-357-358-359-360-361-362-363-364-365-366-367-368-369-370-371-372-373-374-375-376-377-378-379-380-381-382-383-384-385-386-387-388-389-390-391-392-393-394-395-396-397-398-399-400-401-402-403-404-405-406-407-408-409-410-411-412-413-414-415-416-417-418-419-420-421-422-423-424-425-426-427-428-429-430-431-432-433-434-435-436-437-438-439-440-441-442-443-444-445-446-447-448-449-450-451-452-453-454-455-456-457-458-459-460-461-462-463-464-465-466-467-468-469-470-471-472-473-474-475-476-477-478-479-480-481-482-483-484-485-486-487-488-489-490-491-492-493-494-495-496-497-498-499-500-501-502-503-504-505-506-507-508-509-510-511-512-513-514-515-516-517-518-519-520-521-522-523-524-525-526-527-528-529-530-531-532-533-534-535-536-537-538-539-540-541-542-543-544-545-546-547-548-549-550-551-552-553-554-555-556-557-558-559-560-561-562-563-564-565-566-567-568-569-570-571-572-573-574-575-576-577-578-579-580-581-582-583-584-585-586-587-588-589-590-591-592-593-594-595-596-597-598-599-600-601-602-603-604-605-606-607-608-609-610-611-612-613-614-615-616-617-618-619-620-621-622-623-624-625-626-627-628-629-630-631-632-633-634-635-636-637-638-639-640-641-642-643-644-645-646-647-648-649-650-651-652-653-654-655-656-657-658-659-660-661-662-663-664-665-666-667-668-669-670-671-672-673-674-675-676-677-678-679-680-681-682-683-684-685-686-687-688-689-690-691-692-693-694-695-696-697-698-699-700-701-702-703-704-705-706-707-708-709-710-711-712-713-714-715-716-717-718-719-720-721-722-723-724-725-726-727-728-729-730-731-732-733-734-735-736-737-738-739-740-741-742-743-744-745-746-747-748-749-750-751-752-753-754-755-756-757-758-759-760-761-762-763-764-765-766-767-768-769-770-771-772-773-774-775-776-777-778-779-780-781-782-783-784-785-786-787-788-789-790-791-792-793-794-795-796-797-798-799-800-801-802-803-804-805-806-807-808-809-810-811-812-813-814-815-816-817-818-819-820-821-822-823-824-825-826-827-828-829-830-831-832-833-834-835-836-837-838-839-840-841-842-843-844-845-846-847-848-849-850-851-852-853-854-855-856-857-858-859-860-861-862-863-864-865-866-867-868-869-870-871-872-873-874-875-876-877-878-879-880-881-882-883-884-885-886-887-888-889-890-891-892-893-894-895-896-897-898-899-900-901-902-903-904-905-906-907-908-909-910-911-912-913-914-915-916-917-918-919-920-921-922-923-924-925-926-927-928-929-930-931-932-933-934-935-936-937-938-939-940-941-942-943-944-945-946-947-948-949-950-951-952-953-954-955-956-957-958-959-960-961-962-963-964-965-966-967-968-969-970-971-972-973-974-975-976-977-978-979-980-981-982-983-984-985-986-987-988-989-990-991-992-993-994-995-996-997-998-999-1000-1001-1002-1003-1004-1005-1006-1007-1008-1009-1010-1011-1012-1013-1014-1015-1016-1017-1018-1019-1020-1021-1022-1023-1024-1025-1026-1027-1028-1029-1030-1031-1032-1033-1034-1035-1036-1037-1038-1039-1040-1041-1042-1043-

communi

RWTH AACHEN
UNIVERSITY

Repositc

Country

The Univ

UB Disc

Reposito

Country
.....

PubDe

Subject(s)

Provider
...

Centre (N

Research Research

Repositio

Country

FDAT is a
primarily

FAIR Data

1100-1101

Subject/

Reposito

Provider

Country

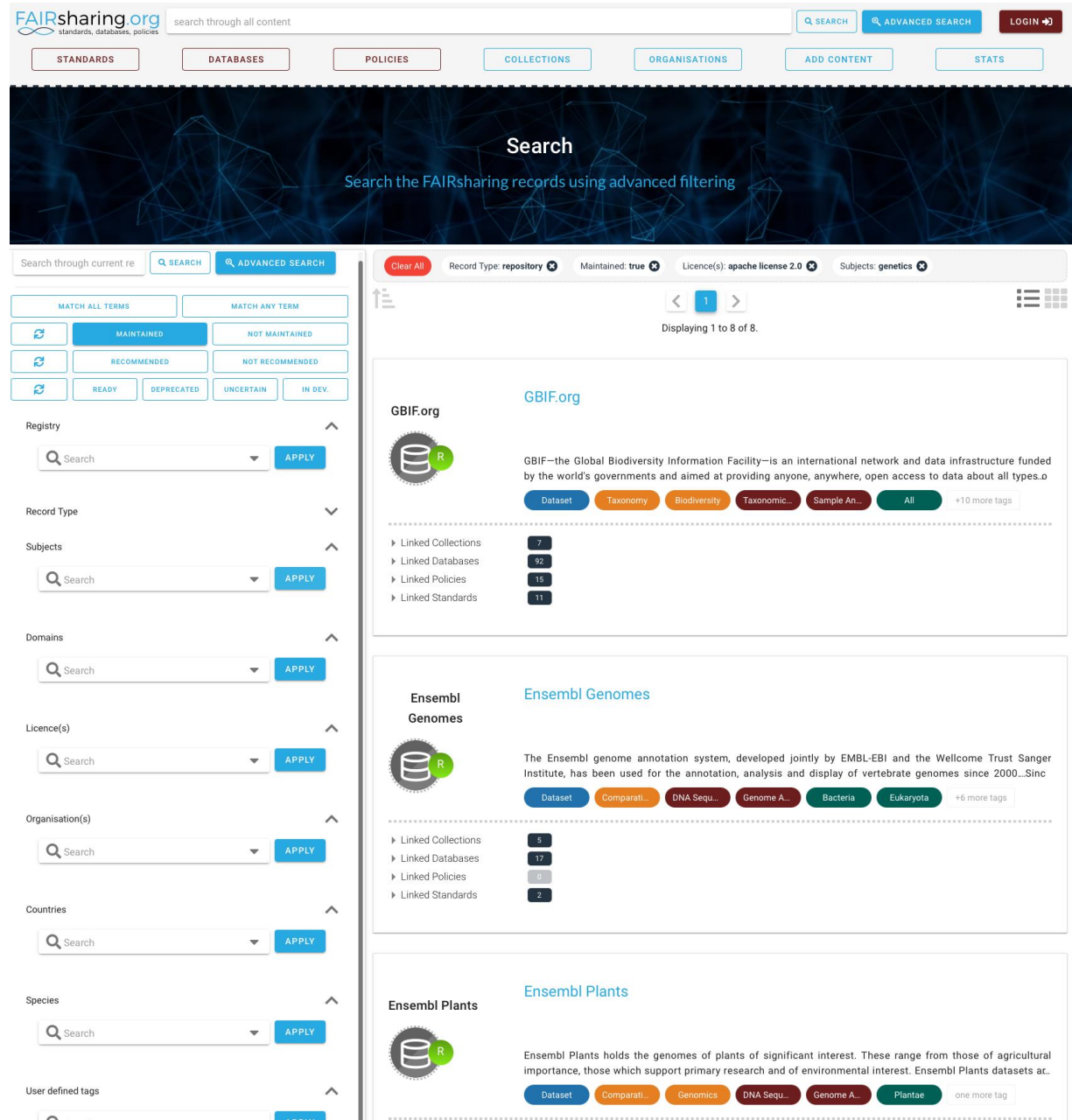
UltraViolet

provides
including

1

ita

FAIRsharing



The screenshot displays the FAIRsharing.org website interface. At the top, there is a navigation bar with links for STANDARDS, DATABASES, POLICIES, COLLECTIONS, ORGANISATIONS, ADD CONTENT, and STATS. A search bar is also present. Below the navigation bar, a large banner reads "Search" and "Search the FAIRsharing records using advanced filtering".

The left sidebar contains various search filters:

- Search through current re
- MATCH ALL TERMS / MATCH ANY TERM
- MAINTAINED / NOT MAINTAINED
- RECOMMENDED / NOT RECOMMENDED
- READY / DEPRECATED / UNCERTAIN / IN DEV.
- Registry (Search, APPLY)
- Record Type (Search, APPLY)
- Subjects (Search, APPLY)
- Domains (Search, APPLY)
- Licence(s) (Search, APPLY)
- Organisation(s) (Search, APPLY)
- Countries (Search, APPLY)
- Species (Search, APPLY)
- User defined tags (Search, APPLY)

The main content area shows search results for "GBIF.org", "Ensembl Genomes", and "Ensembl Plants". Each result includes a description, tags, and a list of linked collections, databases, policies, and standards.

GBIF.org

GBIF—the Global Biodiversity Information Facility—is an international network and data infrastructure funded by the world's governments and aimed at providing anyone, anywhere, open access to data about all types.o

Dataset Taxonomy Biodiversity Taxonomic... Sample An... All +10 more tags

Linked Collections: 7
Linked Databases: 92
Linked Policies: 15
Linked Standards: 11

Ensembl Genomes

The Ensembl genome annotation system, developed jointly by EMBL-EBI and the Wellcome Trust Sanger Institute, has been used for the annotation, analysis and display of vertebrate genomes since 2000...Sinc

Dataset Comparati... DNA Sequ... Genome A... Bacteria Eukaryota +6 more tags

Linked Collections: 5
Linked Databases: 17
Linked Policies: 8
Linked Standards: 2

Ensembl Plants

Ensembl Plants holds the genomes of plants of significant interest. These range from those of agricultural importance, those which support primary research and of environmental interest. Ensembl Plants datasets ar...

Dataset Comparati... Genomics DNA Sequ... Genome A... Plantae one more tag

ERC

Repository name	Repository short name (if available)	Repository URL	Is the repository trusted?	Trusted - Certified repository	Trusted - Community Endorsement	Trusted - Basic Features	Exemplary Readiness Level	Essential Readiness Level	Close to Essential Readiness Level	Data Inconclusive	Coverage	ERC Panels	Content type	Specific Literature type	Access to repository content	Access to data upload	Presence in OpenAIRE	Certification	Community Endorsement	International y recognised	Public Policy for Preservation, Curation and Security of the content	Policy URL
4TU.ResearchData		https://data.4tu.nl/	Yes	Yes	No	Yes	No	No	Yes	No	Disciplinary;	PE1 Mathematical Data; Software/Code			No barriers	Country specific	Yes, my repository Core Trust Seal	Yes	Yes	No	Yes	https://data.4tu.nl/
ARCHE - A Research Data Repository	ARCHE	https://arche.eoscportal.org/	Yes	Yes	Yes	Yes	No	No	No	Yes	Disciplinary;	SH4 The Human Data; Software/Code; Other			No barriers	Community/D	Yes, my repository Core Trust Seal	Yes	Yes	Yes	Yes	https://arche.eoscportal.org/
Biological and BioStudies	BCO-DMO	https://www.bco-dmo.org/	Yes	Yes	Yes	Yes	No	No	Yes	No	Disciplinary;	PE10 Earth Systems Data; Software/Code; Other			No barriers	Funder specific	No, my repository Core Trust Seal	Yes	Yes	Yes	Yes	https://www.bco-dmo.org/
DANS Data Station Social Sciences		https://www.ebi.ac.uk/biostudies/	Yes	No	Yes	Yes	No	No	Yes	No	Disciplinary;	LS1 Molecules Data	Reports, projects		No barriers	Community/D	Not yet/Not sure	No certification	Yes	Yes	Yes	https://www.ebi.ac.uk/biostudies/
		https://ssh.dans.knaw.nl/	Yes	No	Yes	Yes	No	No	Yes	No	Disciplinary;	SH1 Individual Data			No barriers	Country specific	Yes, my repository Core Trust Seal	No certification	Yes	Yes	Yes	https://dans.knaw.nl/
Finnish Social Science Data	FSD	https://www.fsd.fi/	Yes	Yes	Yes	No	No	No	Yes	No	Disciplinary;	SH1 Individual Data			No barriers	Other	Yes, my repository Core Trust Seal	Yes	Yes	Yes	Not Applicable	
GFZ Data Services		https://dataset.gfz.net/	Yes	No	Yes	Yes	No	No	Yes	No	Disciplinary;	PE10 Earth Systems Data; Software/Code			No barriers	Community/D	Not yet/Not sure	No certification	Yes	Yes	Yes	https://dataset.gfz.net/
Lithuanian Data	LIDA	https://lida.dia.lt/	Yes	No	Yes	No	No	No	Yes	No	Disciplinary;	SH1 Individual Data; Software/Code			Authentication	Other	Yes, my repository Core Trust Seal	No certification	Yes	Yes	Not Applicable	
PANGAEA - Data for Global Earth Science	PANGAEA	https://www.pangaea.de/	Yes	Yes	Yes	Yes	No	No	Yes	No	Disciplinary;	PE10 Earth Systems Data			No barriers	Community/D	Yes, my repository Core Trust Seal	Yes	Yes	Yes	Yes	https://www.pangaea.de/
Protein Data Bank	PDB	https://pdbe.org/	Yes	Yes	Yes	Yes	No	No	No	Yes	Disciplinary;	LS1 Molecules Data			No barriers	Community/D	Yes, my repository Core Trust Seal	Yes	Yes	Yes	Yes	https://www.pdb.org/
Social Data Network	SoDaNet	https://sodanet.nl/	Yes	No	Yes	Yes	No	No	Yes	No	Disciplinary;	SH1 Individual Data			No barriers	Community/D	No, my repository Core Trust Seal	No certification	Yes	Yes	Yes	https://sodanet.nl/
The Austrian Social Science Data	AUSSDA	https://data.aussda.at/	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Disciplinary;	SH1 Individual Data; Software/Code			No barriers	Community/D	Yes, my repository Core Trust Seal	Yes	Yes	Yes	Yes	https://aussda.at/
The Tromsø Research Centre	TROLLING	https://trolling.no/	Yes	No	Yes	Yes	No	No	Yes	No	Disciplinary;	PE6 Computer Data; Software/Code			No barriers	Community/D	Yes, my repository Core Trust Seal	No certification	Yes	Yes	Yes	https://site.uit.no/
UK Polar Data	UK PDC	https://www.ukpdc.org/	Yes	Yes	Yes	No	No	No	No	Yes	Disciplinary;	PE2 Fundamental Data			No barriers	Community/D	Not yet/Not sure	Core Trust Seal	Yes	Yes	Not Applicable	

Metadata contains "License" field	PID Assignment	Machine-actionable	Standardized Metadata	Metadata Standard	Linked resources	Linked resources PID(s)	CC0 or PD or equivalent Metadata	Author(s)	Description	Title	Date of publication/deposit	Venue of publication/deposit	Embargo	Metadata allow reference to "Horizon Europe"	Grant Project name	Grant project acronym	Grant project number	Record PID	Author(s) PID	Organisation	Grant PID
Yes	Yes	Yes	Yes	Dublin Core; DDI	Yes, through a	Yes, through a	Yes	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	No	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a
Yes	Yes	Yes	Yes	We have a dedicated schema.org	Yes, through a	Yes, through a	Yes	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a
Yes	Yes	Yes	Yes	ISO19115-2, DDI	Yes, through a	Yes, through a	Yes	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a
Yes	Yes	Yes	Yes	DataCite	Yes, through a	Yes, through a	Yes	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	No	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a
Yes	Yes	Yes	Yes	DDI	Yes, through a	Yes, through a	Yes	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	No	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a
Yes	Yes	Yes	Yes	DataCite, ISO19115-2, DDI	Yes, through a	Yes, through a	Yes	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a
Yes	Yes	Yes	Yes	Dublin Core, DDI	Yes, through a	Yes, through a	Yes	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a
Yes	Yes	Yes	Yes	Dublin Core, PDBx/mmCIF	Yes, through a	Yes, through a	Yes	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a
Yes	Yes	Yes	Yes	DDI	Yes, through a	Yes, through a	Yes	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a
Yes	Yes	Yes	Yes	CESSDA Metadata	Yes, through a	Yes, through a	Yes	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a
Yes	Yes	Yes	Yes	See the section	Yes, through a	Yes, through a	Yes	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a
Yes	Yes	Yes	Yes	ISO 19115	Yes, through a	Yes, through a	Yes	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a	Yes, through a

Deposits

Good practice

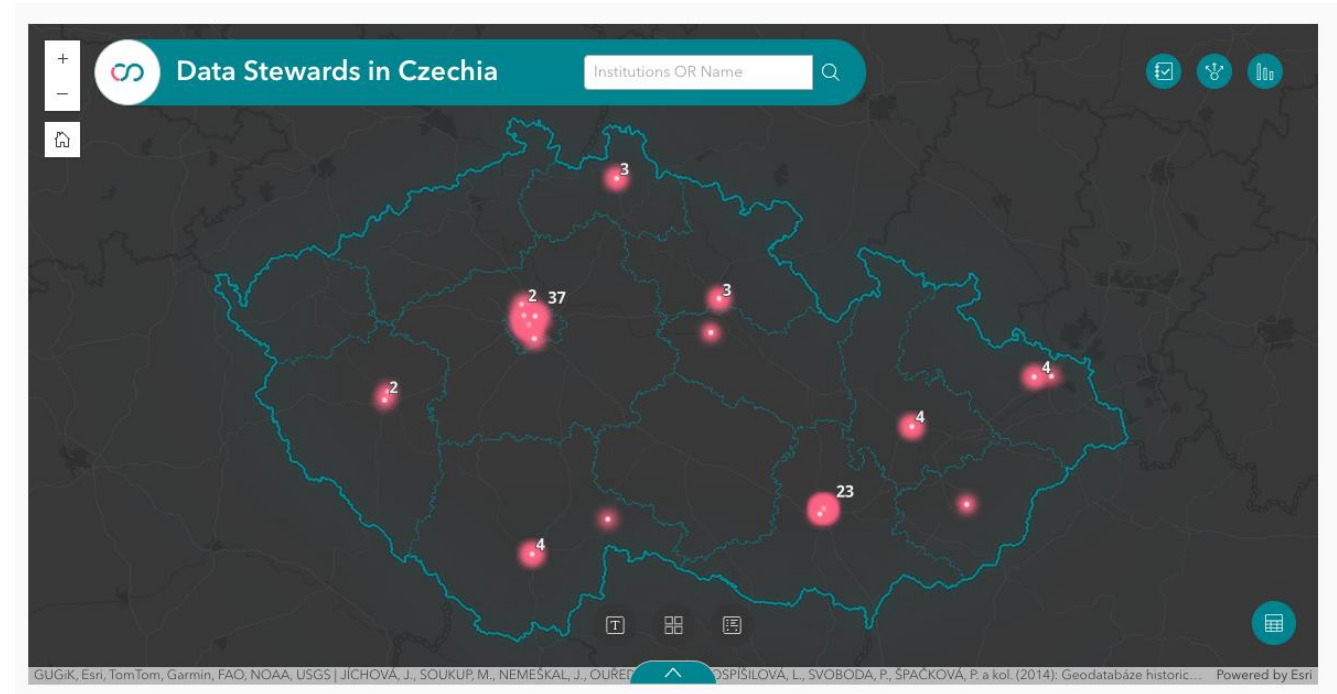
- Title – descriptive & concise
- Metadata – common (e.g., DataCite) & field-specific
- Formats – preferred
- Organization – clear, shallow, logical, well-described, no RAR
- File names – no special characters, consistent, unique, YYYYMMDD
- ReadMe file
 - title, DOI, last update, contact, license, methods, SW, formats, units, naming
 - TXT or PDF/A
 - UTF-8

Common issues

- Publishing
 - Unsuitable data (IP, sensitivity ...)
 - Non-preferred formats
 - Without ReadMe
- Embargo/controlled access
 - not using one when suitable
 - using when unsuitable
- Missing/incorrect metadata
 - PIDs
 - *Creator* vs. *Contributor*
 - Grant dedication
 - License
 - Qualified references
- Not linking dataset in article's data availability statement

Where to get support?

- Your institutional
 - Data Stewards
 - Librarians
- [NTK](#) – free consultations
- Community
 - [Map of Data Stewards](#)
 - [Discord](#)



What to take home?

- Motivation → requirements → repository selection
- Consider places, where you already retrieve data from
- Use the aggregators or your funders for overview
- If in doubt, reach out for support – the community is here for this

Questions?

jan.valis@techlib.cz 



Co-funded by
the European Union

