

TOOLS EXCLUDED FROM THE ANALYSIS

This document provides lists of the tools that were excluded by each of the exclusion criteria during the systematic search of our paper “A Survey of Data Catalog Tools” (<DOI will be inserted here>).

Exclusion Criterion E1: No Support or Further Information Available

The 85 tools that were excluded by the first exclusion criterion (E1: no support or further information available) are listed below.

For each tool, it is listed whether:

- no further information besides the initial mentioning is available (✓)

or whether

- it was removed due to being not maintained since 2018 (if applicable the year of the last update or commit is listed),
- the source code repository is archived or
- the tool is discontinued by its vendor.

A “-” denotes that this part of the criterion is not applicable, since the other part of the criterion led to exclusion.

The table below is sorted in ascending order by tool name.

Tool name	Vendor / Developer	URL	No further information available	No update since 2018 / Repository is archived / Discontinued
A metadata catalog for organization and systemization of fusion simulation data	-	https://www.researchgate.net/publication/257318004_A_metadata_catalog_for_organization_and_systemization_of_fusion_simulation_data	-	2014
A Metadata-Driven Tool for FAIR Data Production in Citizen Science Platforms	-	https://www.researchgate.net/publication/361657596_A_Metadata-Driven_Tool_for_FAIR_Data_Production_in_Citizen_Science_Platforms	-	2008
Adaptive Metadata Manager	Adaptive	https://www.adaptive.com/products/metadata-management	✓	-
Aggua	Aggua	https://www.aggua.io/	-	2011
AirScope	-	https://ieeexplore.ieee.org/document/6528653	-	2013
Alex Solutions	Alex	https://alexolutions.com/product/augmented-data-catalog/	✓	-
AMGA Metadata Catalog	-	https://www.researchgate.net/publication/1959284_Distributed_Metadata_with_the_AMGA_Metadata_Catalog	-	2006
Azure Data Catalog	Microsoft Azure Data Catalog	https://azure.microsoft.com/de-de/products/data-catalog/	✓	-
CAROS	-	https://ieeexplore.ieee.org/document/6744740	-	2007
Catalog Manager for IMS	BMC	https://www.bmc.com/documents/datasheets/catalog-manager-for-ims.html	✓	-
CEOS WGISS Integrated Catalogue (CWIC)	-	https://www.researchgate.net/publication/272549074_Federated_Catalogue_for_Discovering_Earth_Observation_Data	-	2009
Chemomentum	-	https://link.springer.com/chapter/10.1007/978-3-642-00955-6_11	✓	-
Data Catalog Project	-	https://www.sciencedirect.com/science/article/abs/pii/S0920379616303398?via%3Dihub	-	2008
Data catalog project—A browsable, searchable, metadata system	-	https://www.researchgate.net/publication/303808883_Data_catalog_project-A_browsable_searchable_metadata_system	✓	-
Data Ecosystem Management Suite	Global Ids	-	-	only support available, not on main homepage
dscribe	dscribe	-	✓	-
eGenVar data management system	-	https://academic.oup.com/database/article/doi/10.1093/database/bau027/2634026	-	2018
ELDER	-	https://github.com/elderonline/ELDER	✓	-
EMR	-	https://link.springer.com/chapter/10.1007/978-3-540-76888-3_86	-	2012
Extensible USN Metadata Management System	-	Extensible USN Metadata Management System IEEE Conference Publication IEEE Xplore	-	2009

Tool name	Vendor / Developer	URL	No further information available	No update since 2018 / Repository is archived / Discontinued
Field33 Platform	Field33	https://field33.com/	✓	-
GEMMS	-	https://ceur-ws.org/Vol-1612/paper17.pdf	✓	-
GEON catalog	-	https://www.researchgate.net/publication/252733599_Federating_Metadata_Catalogs	✓	-
GI-go	-	https://www.researchgate.net/publication/252622982_Geo-Information_Catalog_Services_Interoperability_an_Experimented_Tool	-	2017
Gorge	-	https://github.com/ZdogMcSuperfly/Gorge	✓	-
GraphMeta	-	https://ieeexplore.ieee.org/document/8580412	-	2012
Infoworks Autonomous Data Engine	Infoworks	https://go.infoworks.io/hubfs/Infoworks%20Brochures/Infoworks%20BI%20&%20Analytics%20Workflow%20Automation%20v%20Jan%2018.pdf	✓	-
Intel Machine Learning Scaling Library	Intel	https://github.com/intel/MLSL	✓	-
IQ Metadata Manager	InQuisient	https://inquisient.com/resource/iq-metadata-manager/	✓	-
Itematix	-	-	✓	-
Junar	-	-	✓	-
Loris-bases Metadata Management System	-	https://ieeexplore.ieee.org/document/6005451	-	2012
LSST Data catalog	-	-	-	2018
madxxx_catalog_client	-	https://github.com/eoss-cloud/madxxx_catalog_client	✓	-
Mercury	-	https://link.springer.com/article/10.1007/s12145-010-0050-7	-	2018
Meta-fy	-	https://projects.cs.uct.ac.za/honsproj/cgi-bin/view/2014/benson_ferguson.zip/downloads/Michael_Ferguson%20Downloads/HONOURS_PROJECT_REPORT_FINAL_MichaelFerguson_29_10_2014.pdf	-	2006
MetaCube	-	https://ieeexplore.ieee.org/document/7751658	✓	-
Metadata Catalogs with Semantic Representations	-	https://www.researchgate.net/publication/220919041_Metadata_Catalogs_with_Semantic_Representations	-	2009
Metadata Management and Harvesting Sytsem in Smart Open Data	-	https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9728796	-	2013
Metadata Management System for Data Warehouses	-	https://ieeexplore.ieee.org/document/5690320	-	2009
Metadata Management System for file system	-	https://ieeexplore.ieee.org/document/6375575	-	2008
metadata management system of linkage service of scientific data and scientific literature	-	https://www.matec-conferences.org/articles/mateconf/pdf/2018/105/mateconf_iswso2018_03026.pdf	-	2014
Metadata Management Tool to control FML	-	https://link.springer.com/chapter/10.1007/978-3-031-23298-5_1	-	2018
Metadata Management Web Tool	-	https://github.com/NYCPlanning/edm-metadata-builder	-	2010
Metadata Management Web Tool	-	https://github.com/amolivani/Metadata	-	2016
metadata-server	-	https://github.com/usgin/metadata-server	✓	-
MetaShare	-	https://ieeexplore.ieee.org/document/7344372	-	2007
MetaStore	-	https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7840956	✓	-
Midas	-	https://link.springer.com/chapter/10.1007/978-3-030-33752-0_9	-	2008
MMS	-	https://ieeexplore.ieee.org/document/4221833	-	2010
MMS for cultural data	-	https://iopscience.iop.org/article/10.1088/1742-6596/1201/1/012037/meta	✓	
MODIS	-	https://ieeexplore.ieee.org/document/4662881	-	2014
Montra	-	https://www.sciencedirect.com/science/article/pii/S0169260717312725	-	2009
Multimedia Metadata Management System	-	https://ieeexplore.ieee.org/document/5367703	-	2018
myLEAD	-	Personal Workspace for Large-Scale Data-Driven Computational Experiment IEEE Conference Publication IEEE Xplore	-	2009

Tool name	Vendor / Developer	URL	No further information available	No update since 2018 / Repository is archived / Discontinued
NebraskaMAP	-	https://www.researchgate.net/publication/280073847_GEOSPATIAL_DATA_PORTALS_LIBRARIANS_ADD_EXPERTISE_IN_THE_DEVELOPMENT_OF_GIS_METADATA_CATALOGS	-	2016
ONEMercury	-	https://link.springer.com/article/10.1007/s00799-015-0146-2#Sec29	✓	-
Prototype for geological applications	-	https://ieeexplore.ieee.org/document/6626058	-	2006
Railway Metadata Management System	-	https://ieeexplore.ieee.org/document/8593216?denied=	-	2016
RCMMS - Distributed Metadata Management System for Data Grid	-	https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4287816	-	2016
Remote sensing image distribute system	-	https://ieeexplore.ieee.org/document/1525226	-	2006
SAP Data Hub	SAP	https://help.sap.com/docs/SAP_DATA_HUB?locale=en-US	-	GitHub repo is archived
SMICC	-	https://ieeexplore.ieee.org/document/5318903	-	2005
Socrata	-	-	-	2009
Spectrum Discovery	Precisely	https://support.precisely.com/products/spectrum-discovery/	✓	-
tame	-	https://github.com/meson800/tame	✓	-
thoth	-	https://github.com/DoppleDankster/thoth	✓	-
Tree Schema Data Catalog	Tree Schema	https://treeschema.com/	✓	-
trocco	primeNumber	https://trocco.io/lp/index.html	✓	-
TV-Anytime metadata management system	-	Metadata Management Based on Set-Top Box in Digital Broadcasting Environments IEEE Conference Publication IEEE Xplore	-	2010
unifi data catalog	unifi software	https://aws.amazon.com/marketplace/pp/prodview-dc4yyvgbvln4	-	2007
Waterline Data Catalog	Waterline	https://www.waterlinedata.com/data-catalog-details/	✓	-
WebSphere Metadata Server	IBM	https://www.ibm.com/docs/en/iis/8.1?topic=services-websphere-metadata-server	-	2005
-	-	https://dl.acm.org/doi/10.1145/2882903.2899389	✓	-
-	-	https://dl.acm.org/doi/10.1145/1096536.1096552	✓	-
-	-	https://dl.acm.org/doi/10.1145/3472163.3472273	✓	-
-	-	https://ieeexplore.ieee.org/document/5289122	-	2011
-	-	https://ieeexplore.ieee.org/document/5231081	✓	-
-	-	https://ieeexplore.ieee.org/document/6210386	-	2013
-	-	https://www.researchgate.net/publication/271627394_Intelligent_polar_cyberinfrastructure_enabling_semantic_search_in_geospatial_metadata_catalogue_to_support_polar_data_discovery	-	2016
-	-	https://www.researchgate.net/publication/367287039_From_simulation_to_dissemination_automation_of_data_and_metadata_management	✓	-
-	-	https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6626035	-	2015
-	-	https://www.sciencedirect.com/science/article/pii/S0168169904001620#aep-section-id11	-	2011
-	-	https://link.springer.com/chapter/10.1007/978-3-540-87442-3_135	-	2013
-	-	https://link.springer.com/chapter/10.1007/978-1-4614-1767-5_3	-	2015

Exclusion Criterion E2: Different Purpose tools

The 60 tools that were excluded by the second exclusion criterion (E2: different purpose tools) are listed below.

For each tool, we list what we identified as its main purpose instead of being a data catalog tool.

The table below is sorted in ascending order by tool name.

Tool name	Vendor / Developer	URL	Purpose
Alteryx Connect	Alteryx	https://www.alteryx.com/de/glossary/data-catalog	Data integration
ancoraDocs	ancora	https://ancorasoftware.com/ancoradocs-enterprise/	Data management and quality
Anzo	Cambridge Semantics	https://cambridgesemantics.com/anzo-platform/	Knowledge graph and semantic tool
Appen	Appen	https://appen.com/platform-5/	Other
Augmented Data Fabric Platform	Stratio	https://www.stratio.com/product	Data integration
Binah.AI	Binah.AI	https://www.binah.ai/	Other
Binary Demand	Binary Demand	https://www.binarydemand.com/solutions/data-solutions/	Other
Claravine	Claravine	https://www.claravine.com/	Data management and quality
Cloudera Data Platform	Cloudera	https://www.cloudera.com/products/cloudera-data-platform.html	BI, observability, and analytics
Coginiti	Coginiti	https://www.coginiti.co/	BI, observability, and analytics
Composable DataOps Platform	Composable Analytics	https://composable.ai/dataops-enterprise-platform/	DataOps
Data Dynamics StorageX	Data Dynamics	https://www.datadynamicsinc.com/product-family/storagex/	Other
Data Harmony	Access Innovations	https://www.accessinn.com/data-harmony/	Knowledge graph and semantic tool
Data Lake Metadata Management System	-	https://github.com/yanzhao-irit/data-lake-metadata-management-system	Data lake and lakehouse
DataHawk	we bridge	https://we-bridge.com/products/datahawk/	Lineage and visualization
Datalogz	Datalogz	https://www.datalogz.io/	BI, observability, and analytics
Denodo Platform	Denodo	https://www.denodo.com/en/denodo-platform/overview	Data integration
Dockit Metadata Manager	Vyapin	https://www.vyapinsoftware.com/products/sharepoint-metadata-management-tool	Data management and quality
DQLabs	DQLabs	https://www.dqlabs.ai/	Data management and quality
Dremio Lakehouse Platform	Dremio	https://www.dremio.com/platform/	Data lake and lakehouse
Flyte	Flyte Project Authors	https://github.com/flyteorg/datacatalog	DataOps
Google Cloud Composer	Google	https://cloud.google.com/composer?hl=de	Other
Google Cloud Dataplex	Google	https://cloud.google.com/dataplex?hl=de	Other
Hygraph	Hygraph	https://hygraph.com/use-cases/inventory-and-catalog-management	Other
IBM Cloud Pak for Data	IBM	https://www.ibm.com/products/cloud-pak-for-data	Data integration
IBM InfoSphere Information Server	IBM	https://www.ibm.com/information-server	Data integration
immuta Data Security Platform	immuta	https://www.immuta.com/	Security
Informatica Axon	Informatica	https://www.informatica.com/de/products/data-quality/informatica-axon.html	Data management and quality
Informatica Data Preperation	Informatica	https://www.informatica.com/products/data-catalog/data-prep.html	Data management and quality
Informatica Data Privacy Management	Informatica	https://www.informatica.com/se/products/data-security/data-privacy-management.html	Security
Invenio	-	https://inveniosoftware.org/	Data management and quality
Inventa	1touch	https://www.1touch.io/	Security
iomete	iomete	https://iomete.com/#process	Data lake and lakehouse
LinkAhead	IndiScale GmbH	https://www.indiscale.com/de/linkahead-2/	Data management and quality
MANTA Platform	MANTA	https://getmanta.com/	Lineage and visualization
metaphactory	metaphacts	https://metaphacts.com/product	Knowledge graph and semantic tool
Monte Carlo	Monte Carlo	https://www.montecarlodata.com/	BI, observability, and analytics

Tool name	Vendor / Developer	URL	Purpose
MRO3i	Net Results Group	https://netresultsgroup.com/net-results-group-unveils-mro3i-4-0-powering-industrial-collaboration/	Data management and quality
Narrative	Narrative	https://www.narrative.io/	Data management and quality
Open Semantic Search	-	https://opensemanticsearch.org	Data integration
OpenText File Intelligence	OpenText	https://www.opentext.de/produkte-und-loesungen/produkte/enterprise-content-management/file-intelligence	Lineage and visualization
Oracle Data Relationship Management	Oracle	https://www.oracle.com/de/performance-management/data-relationship-management/	Knowledge graph and semantic tool
PoolParty Semantic Suite	Semantic Web Company	https://de.poolparty.biz/	Knowledge graph and semantic tool
Pop SQL Data Catalog	Pop SQL	https://popsql.com/product/data-catalog	Data management and quality
QuiltData	-	https://quiltdata.com/ ; https://github.com/quiltdata/quilt	Other
Raynet Unified Data Platform	Raynet	https://raynet-inc.com/unified-data-management/unified-data-platform/	Data management and quality
Reltio Connected Data Platform	Reltio	https://www.reltio.com/connected-data-platform/	Data management and quality
Sama	Samasource Impact Sourcing	https://www.sama.com/de/#	Other
SAP Data Intelligence	SAP	https://www.sap.com/products/technology-platform/data-intelligence.html	Data integration
SAP PowerDesigner	SAP	https://www.sap.com/products/technology-platform/powerdesigner-data-modeling-tools.html	Other
SAS Viya	SAS Institute	https://www.sas.com/en_nz/software/viya.html	BI, observability, and analytics
Secuvy AI	Secuvy AI	https://secuvy.ai/	Security
Semaphore	MarkLogic	https://www.marklogic.com/product/semaphore/	Data management and quality
Stemma Data Catalog	Stemma	https://www.stemma.ai/	BI, observability, and analytics
Syniti Knowledge Platform	Syniti	https://www.syniti.com/platform/syniti-knowledge-platform/	Data management and quality
Tableau Desktop	Tableau	https://www.tableau.com/de-de/products/desktop	Lineage and visualization
Talend Data Fabric	Talend	https://www.talend.com/ps/data-fabric-demo-request/	Data integration
Tibco Data Virtualization	Tibco	https://www.tibco.com/products/data-virtualization	Data integration
TopBraid Enterprise Data Governance	TopQuadrant	https://www.topquadrant.com/	Knowledge graph and semantic tool
Untrite	Untrite	https://untrite.com/product/	Other

Exclusion Criterion E3: Domain-specific Tools

The 22 tools that were excluded by the third exclusion criterion (E3: domain-specific tools) are listed below.

For each tool, we list for which domain it was created for.

The table below is sorted in ascending order by tool name.

Tool name	Vendor / Developer	URL	Domain
Building an integrated enhanced virtual research environment metadata catalogue	-	https://www.researchgate.net/publication/335883359_Building_an_integrated_enhanced_virtual_research_environment_metadata_catalogue_The_Electronic_Library_Journal_Emerald_Publishing_Limited	Geospatial and environmental
CureSCi MDC	-	https://www.researchgate.net/publication/366211536_CureSCi_Metadata_Catalog-Making_sickle_cell_studies_findable	Biological and medical
Dataverse	-	https://dataverse.org/	Scientific and research
e-Research Distributed Data System (Eddy)	-	https://github.com/monash-merc/eddy	Scientific and research
illuminate Clinical Data Cloud	eClinical Solutions	https://www.eclinicalsol.com/products/	Biological and medical
EMIF	-	https://www.sciencedirect.com/science/article/pii/S138650561830830X	Biological and medical
Empress	-	https://github.com/mlawsonca/empress ; https://dl.acm.org/doi/pdf/10.1145/3523698	Scientific and research
Fedora	-	https://fedora.lyrasis.org/	Cultural heritage
GeoLOD	-	https://www.researchgate.net/publication/350977399_GeoLOD_A_Spatial_Linked_Data_Catalog_and_Recommender	Geospatial and environmental
medna-metadata	-	https://www.researchgate.net/publication/362674377_medna-metadata_an_open-source_data_management_system_for_tracking_environmental_DNA_samples_and_metadata	Geospatial and environmental
MetaCat	-	https://www.researchgate.net/publication/354071981_MetaCat_-_metadata_catalog_for_data_management_systems	Scientific and research
Metadata Management Tool Application	NASA	https://github.com/nasa/mmt	Cultural heritage
Metadata Management Tool for Learning Objects	-	https://vtechworks.lib.vt.edu/bitstream/handle/10919/103649/Okoth_DO_D_2019.pdf?sequence=1&isAllowed=y	Education
Molgenis	-	https://www.molgenis.org/	Scientific and research
Nada	-	https://github.com/ihsn/nada	Scientific and research
PHEMI Health DataLab	PHEMI	https://www.phemi.com/	Biological and medical
SCAR Antarctic Digital Database (ADD) Metadata Toolbox	-	https://github.com/antarctica/scar-add-metadata-toolbox	Geospatial and environmental
SciCat	-	https://scicatproject.github.io/	Scientific and research
The Multiple Sclerosis Data Alliance Catalogue	-	https://www.researchgate.net/publication/357414900_The_Multiple_Sclerosis_Data_Alliance_Catalogue	Biological and medical
VuFind	-	https://vufind.org/vufind/	Cultural heritage
Zenodo	-	https://zenodo.org/	Scientific and research
-	-	https://dl.acm.org/doi/10.1145/3523698 ; https://github.com/mlawsonca/empress	Scientific and research

Exclusion Criterion E4: Tools storing data

The 5 tools that directly store data and thus were excluded by the fourth exclusion criterion (E4: tools storing data) are listed below.

The table below is sorted in ascending order by tool name.

Tool name	Vendor / Developer	URL
Clowder	-	https://dl.acm.org/doi/10.1145/3219104.3219159 ; https://github.com/clowder-framework/clowder
DataHub	-	https://datahub.io
DKAN	-	https://getdkan.org/
iRods	-	https://irods.org/
Slingshot Data Catalog	Slingshot	https://www.slingshotapp.io/product-tour/data-catalog