

DMTC Stats

Karl Benedict

30 March, 2023

Summary Statistics and Visualizations of Content of the Data Management Training Clearinghouse

The document provides a summary of the content of the Data Management Training Clearinghouse that was retrieved¹ from the development server² for the IMLS DMTC.

Growth of the DMTC Collection through Time

Month	Cumulative Resources Added	Cumulative Published Resources
2016-09-01	16	16
2016-11-01	46	46
2016-12-01	47	47
2017-01-01	48	48
2017-02-01	54	54
2017-03-01	55	55
2017-04-01	63	61
2017-05-01	64	62
2017-07-01	72	70
2017-08-01	80	78
2017-10-01	90	88
2017-11-01	94	92
2017-12-01	110	106
2018-01-01	120	116
2018-02-01	155	151
2018-03-01	169	165
2018-04-01	202	197
2018-05-01	214	208
2018-06-01	215	208
2018-07-01	216	208
2018-08-01	295	274
2018-09-01	329	305
2018-10-01	344	320
2018-11-01	362	338
2018-12-01	372	344
2019-01-01	394	365

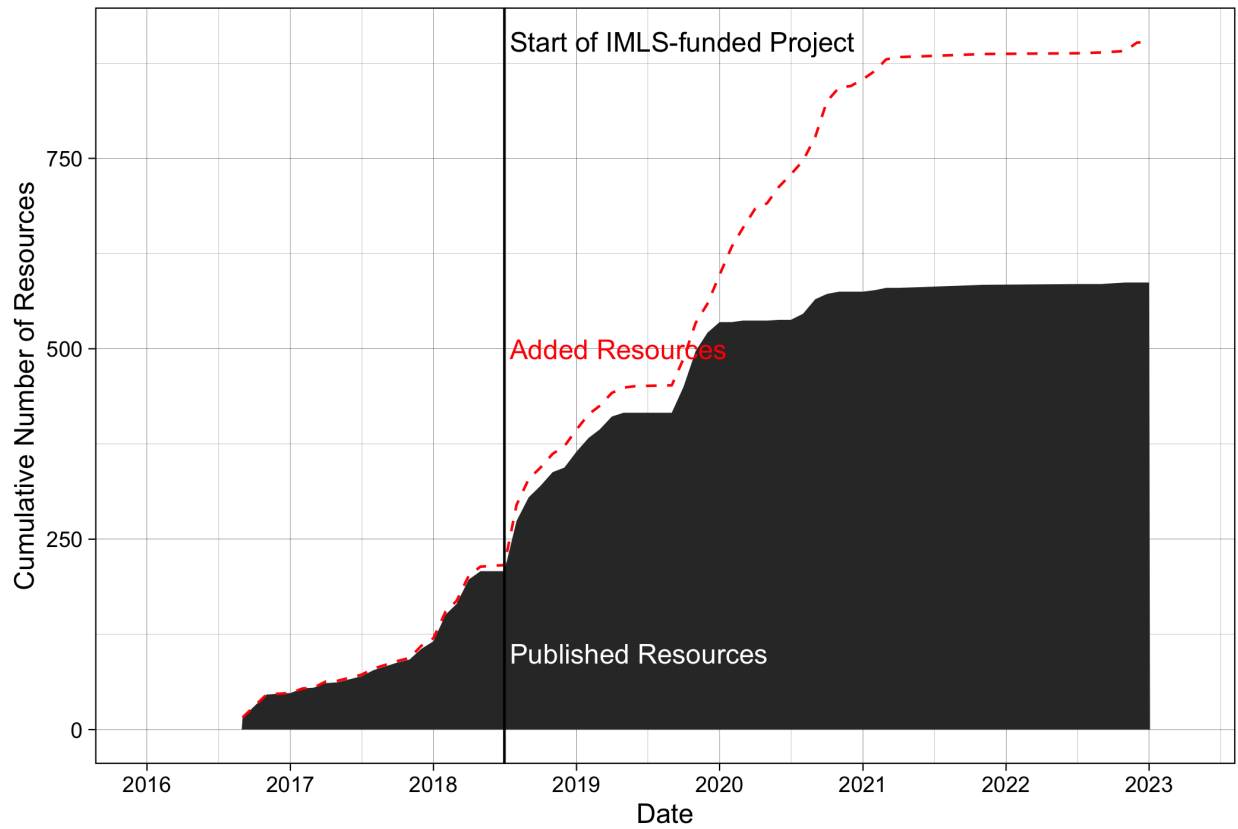
¹Sample request URL: https://esip-dev-02.edacnm.org/api/resources/?facet_limit=0&limit=100&offset=900&api-key=0b8c224a-3c27-11ec-a7b2-9585474babb5 retrieved on 2023-03-30 12:03:30

²<https://esip-dev-02.edacnm.org>

(continued)

Month	Cumulative Resources Added	Cumulative Published Resources
2019-02-01	414	383
2019-03-01	425	394
2019-04-01	442	411
2019-05-01	449	416
2019-06-01	451	416
2019-09-01	452	416
2019-10-01	487	450
2019-11-01	534	497
2019-12-01	559	521
2020-01-01	597	535
2020-02-01	634	535
2020-03-01	659	537
2020-04-01	684	537
2020-05-01	691	537
2020-06-01	713	538
2020-07-01	729	538
2020-08-01	747	546
2020-09-01	778	565
2020-10-01	825	572
2020-11-01	843	575
2020-12-01	845	575
2021-01-01	854	575
2021-02-01	865	577
2021-03-01	880	580
2021-04-01	883	580
2021-06-01	884	581
2021-11-01	887	584
2022-07-01	888	585
2022-09-01	889	585
2022-11-01	891	587
2022-12-01	902	587
2023-01-01	903	587

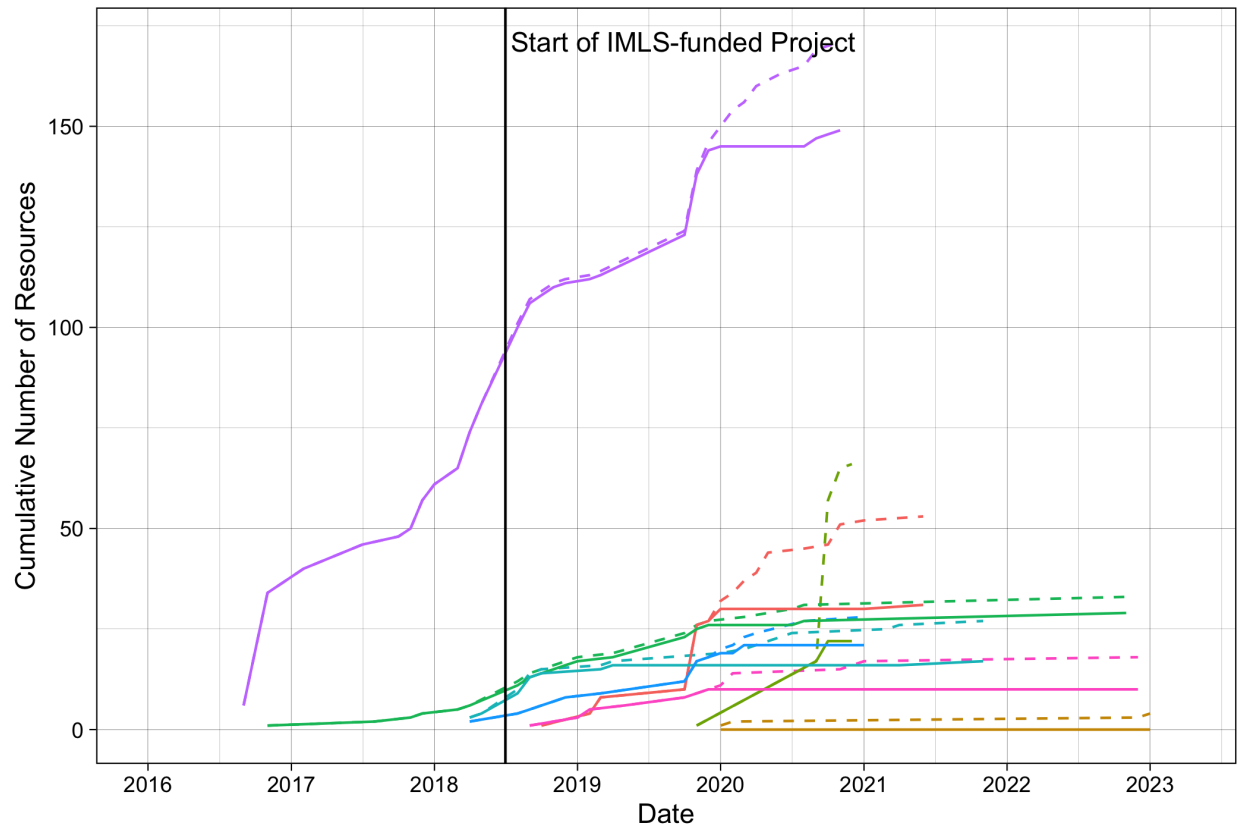
All Added and Published Resources in the DMTC



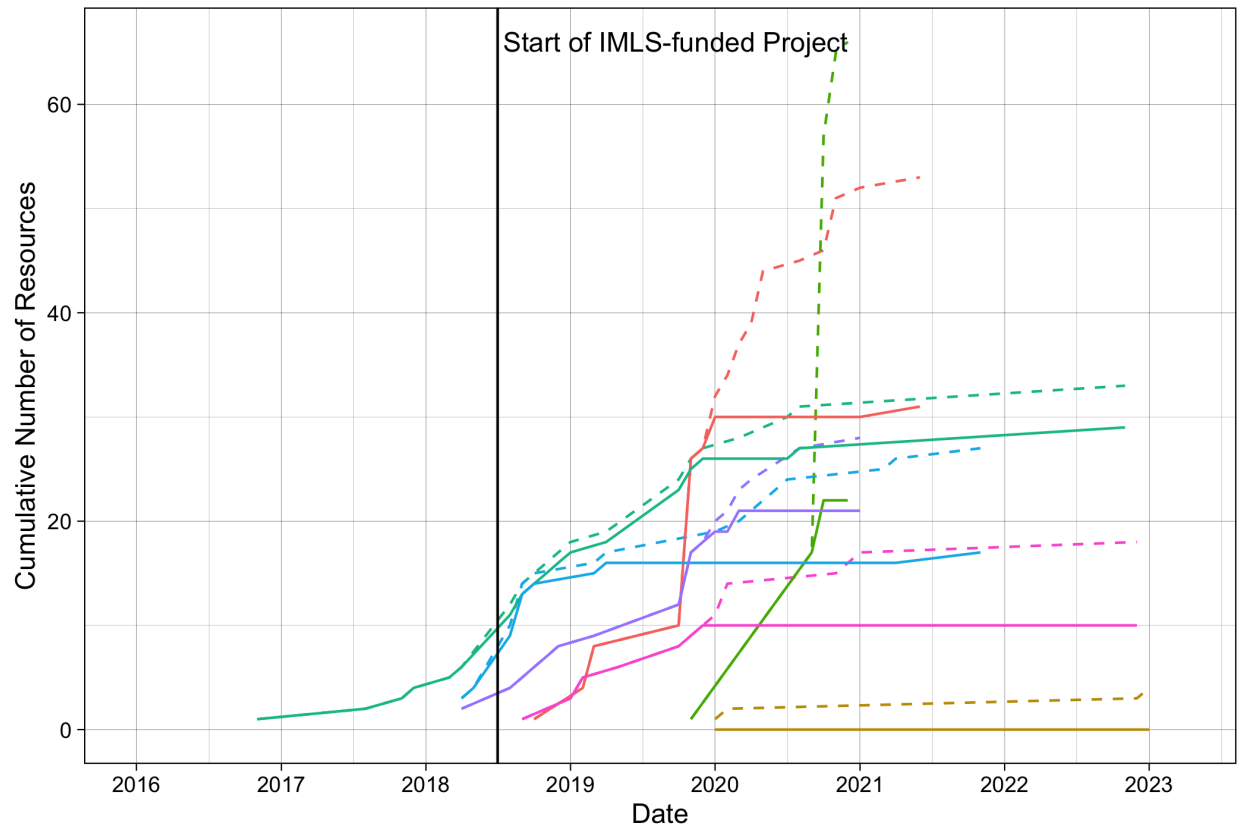
Subject Categories Represented in the DMTC

First Added	Most Recent	Subject Category	Total
2016-09-01	2020-11-01	Physical Sciences and Mathematics	171
2016-11-01	2022-11-01	Engineering	33
2018-04-01	2021-11-01	Life Sciences	27
2018-04-01	2021-01-01	Medicine and Health Sciences	28
2018-09-01	2022-12-01	Social and Behavioral Sciences	18
2018-10-01	2021-06-01	Arts and Humanities	53
2019-11-01	2020-12-01	Education	66
2020-01-01	2023-01-01	character(0)	4

Subject Categories in the DMTC



Subject Categories in the DMTC - Excluding Physical Sciences and Mathematics



Subject Category

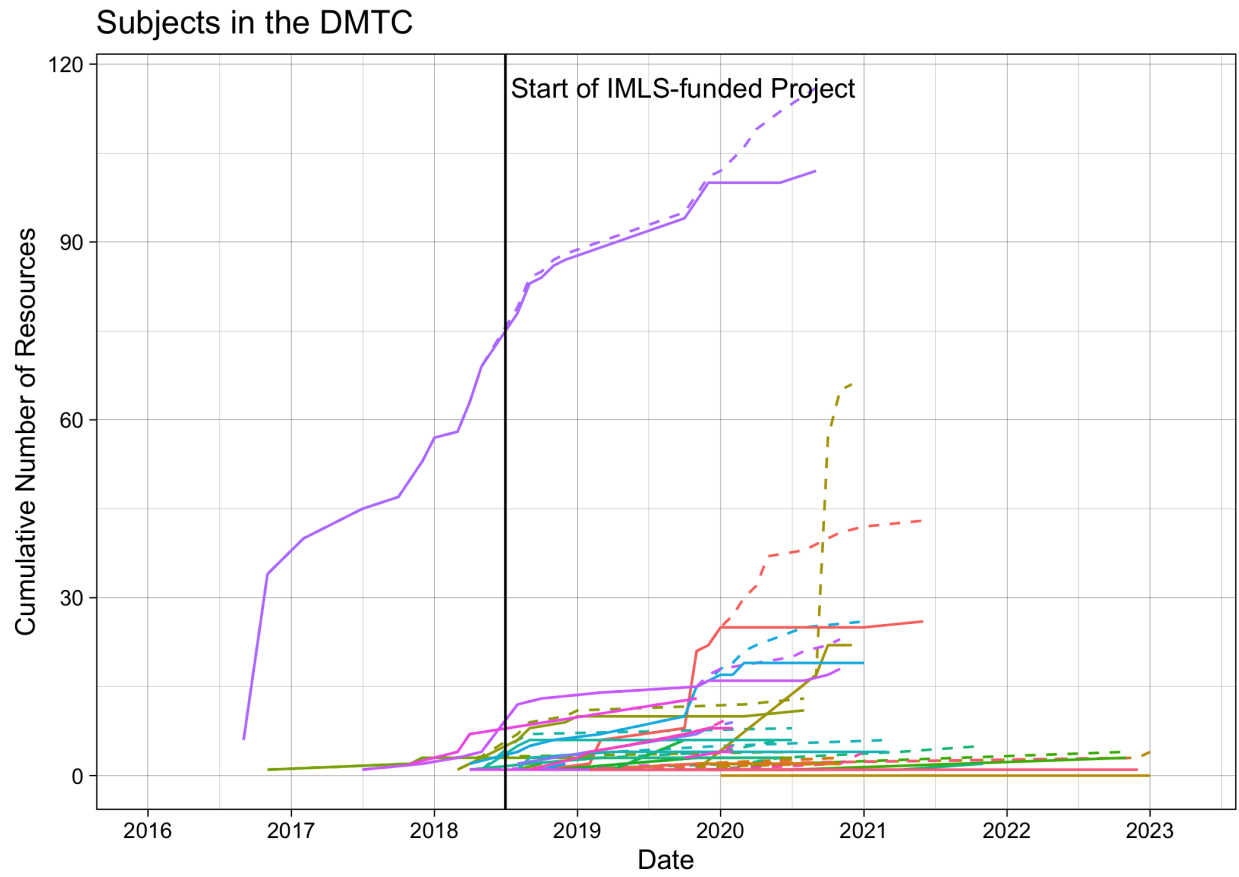
- Arts and Humanities
- character(0)
- Education
- Engineering
- Life Sciences
- Medicine and Health Sciences
- Physical Sciences and Mathematics
- Social and Behavioral Sciences

Detailed Subject Areas Represented in the DMTC

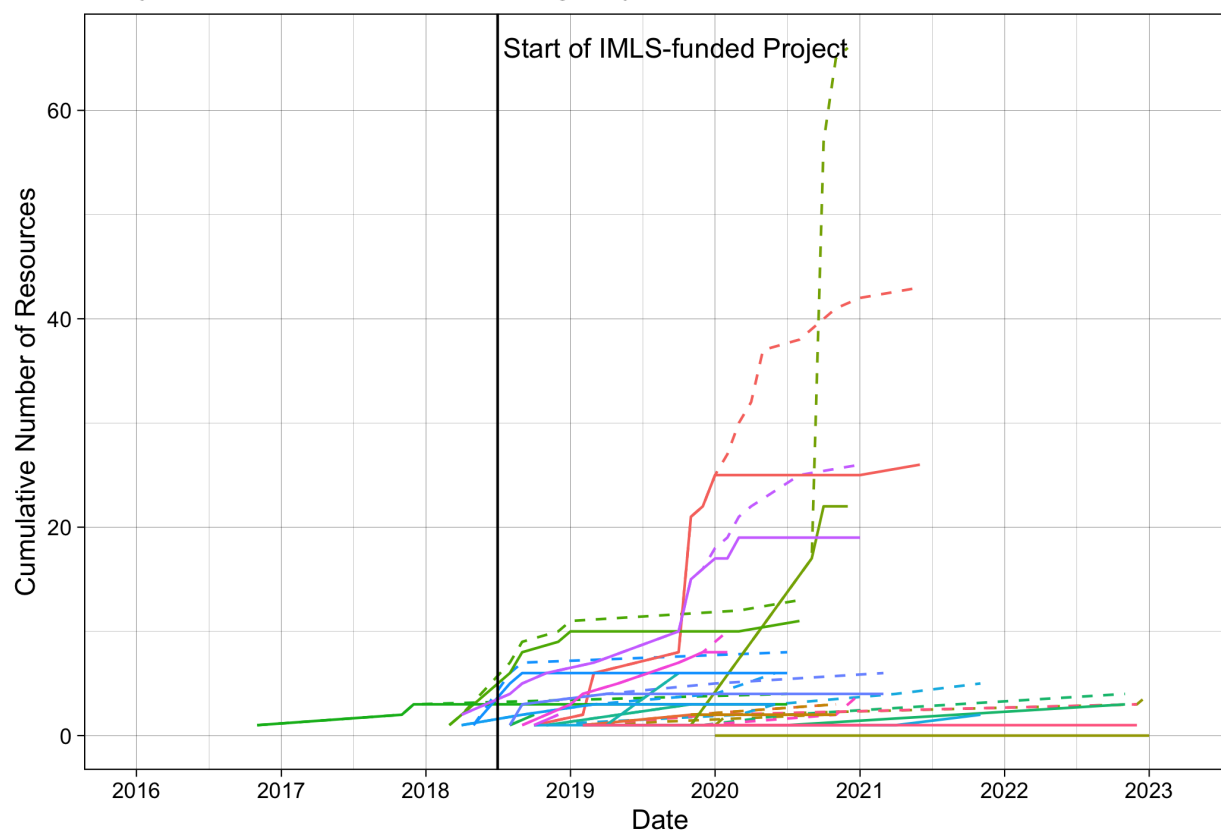
First Added	Most Recent	Subject Category	Total
2016-09-01	2020-09-01	Physical Sciences and Mathematics: Earth Sciences	116
2016-11-01	2020-07-01	Engineering: Aerospace Engineering	4
2017-07-01	2020-11-01	Physical Sciences and Mathematics: Environmental Sciences	23
2017-08-01	2017-08-01	Engineering: Electrical and Computer Engineering	1
2017-11-01	2019-11-01	Physical Sciences and Mathematics: Oceanography and Atmospheric Sciences and Meteorology	13
2018-03-01	2020-08-01	Engineering	13
2018-04-01	2020-06-01	Life Sciences: Agriculture	6
2018-04-01	2018-04-01	Life Sciences: Nutrition	2
2018-04-01	2021-01-01	Medicine and Health Sciences	26
2018-04-01	2020-01-01	Physical Sciences and Mathematics: Mathematics	2

(continued)

First Added	Most Recent	Subject Category	Total
2018-05-01	2020-07-01	Life Sciences: Genetics and Genomics	8
2018-08-01	2018-10-01	Engineering: Biomedical Engineering and Bioengineering	2
2018-08-01	2021-03-01	Life Sciences: Microbiology	6
2018-08-01	2020-02-01	Physical Sciences and Mathematics: Computer Sciences	9
2018-09-01	2018-09-01	Physical Sciences and Mathematics	1
2018-09-01	2020-02-01	Social and Behavioral Sciences	10
2018-10-01	2021-06-01	Arts and Humanities	43
2018-10-01	2021-11-01	Life Sciences	5
2018-10-01	2018-12-01	Medicine and Health Sciences: Biomedical Science	2
2018-11-01	2019-11-01	Engineering: Chemical Engineering	3
2019-02-01	2020-11-01	Arts and Humanities: Digital Humanities	3
2019-02-01	2020-10-01	Arts and Humanities: Other Languages, Societies, and Cultures	2
2019-02-01	2019-02-01	Physical Sciences and Mathematics: Astrophysics and Astronomy	1
2019-02-01	2022-12-01	Social and Behavioral Sciences: Library and Information Science	3
2019-04-01	2019-10-01	Engineering: Computer Engineering	6
2019-10-01	2020-02-01	Physical Sciences and Mathematics: Statistics and Probability	5
2019-11-01	2020-11-01	Arts and Humanities: History of Art, Architecture, and Archaeology	3
2019-11-01	2020-12-01	Education: Science and Mathematics Education	66
2019-12-01	2022-11-01	Engineering: Bioresource and Agricultural Engineering	4
2019-12-01	2019-12-01	Physical Sciences and Mathematics: Chemistry	1
2019-12-01	2021-01-01	Social and Behavioral Sciences: Linguistics	4
2020-01-01	2020-01-01	Arts and Humanities: History	2
2020-01-01	2023-01-01	character(0)	4
2020-02-01	2020-02-01	Social and Behavioral Sciences: Agricultural and Resource Economics	1

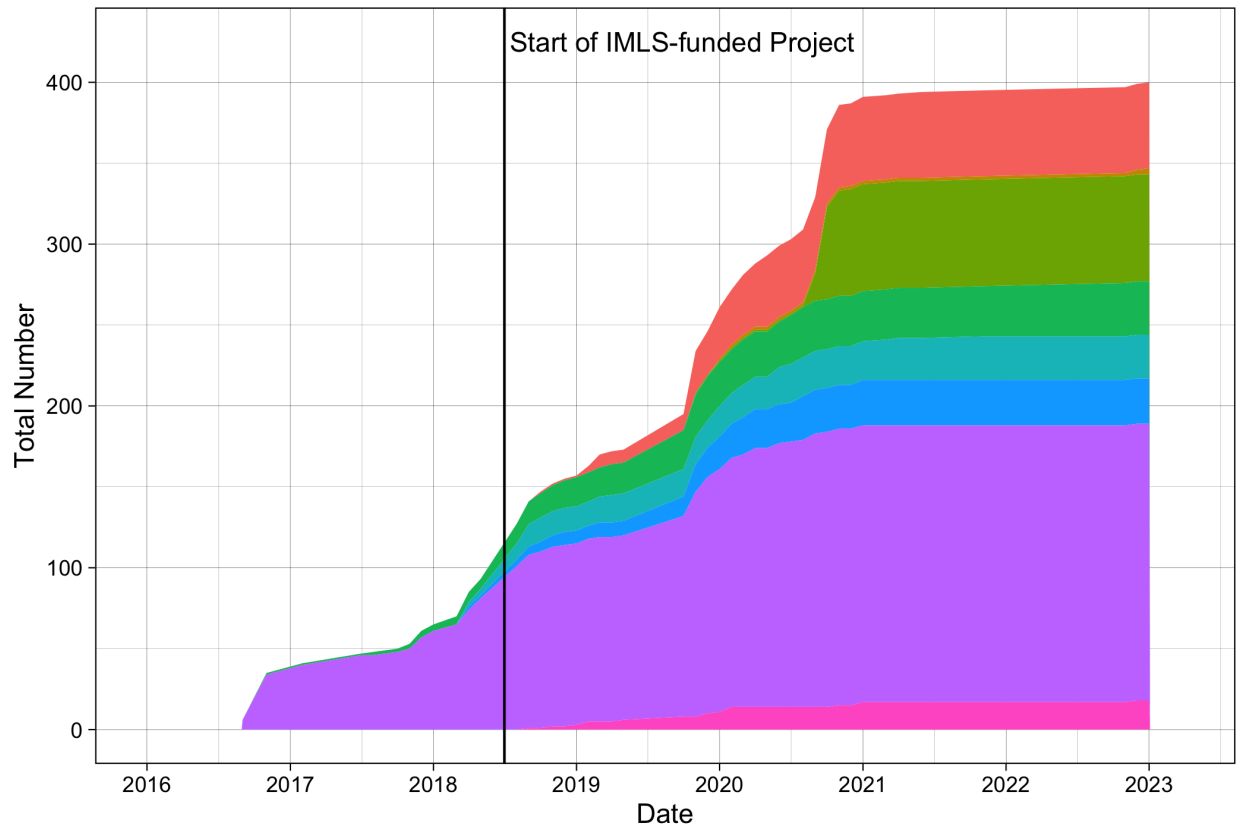


Subjects in the DMTC - Excluding Physical Sciences and Mathematics



Subject Category	
Arts and Humanities	Life Sciences: Microbiology
Arts and Humanities: Digital Humanities	Life Sciences: Nutrition
Arts and Humanities: History	Medicine and Health Sciences
Arts and Humanities: History of Art, Architecture, and Archaeology	Medicine and Health Sciences: Biomedical Science
Arts and Humanities: Other Languages, Societies, and Cultures	Physical Sciences and Mathematics
character(0)	Physical Sciences and Mathematics: Astrophysics and Astronomy
Education: Science and Mathematics Education	Physical Sciences and Mathematics: Chemistry
Engineering	Physical Sciences and Mathematics: Computer Sciences
Engineering: Aerospace Engineering	Physical Sciences and Mathematics: Earth Sciences
Engineering: Biomedical Engineering and Bioengineering	Physical Sciences and Mathematics: Environmental Sciences
Engineering: Bioresource and Agricultural Engineering	Physical Sciences and Mathematics: Mathematics
Engineering: Chemical Engineering	Physical Sciences and Mathematics: Oceanography and Atmospheric Sciences and Meteorology
Engineering: Computer Engineering	Physical Sciences and Mathematics: Statistics and Probability
Engineering: Electrical and Computer Engineering	Social and Behavioral Sciences
Life Sciences	Social and Behavioral Sciences: Linguistics
Life Sciences: Agriculture	Social and Behavioral Sciences: Agricultural and Resource Economics
Life Sciences: Genetics and Genomics	Social and Behavioral Sciences: Library and Information Science

Subject Categories in the DMTC



Subject Category

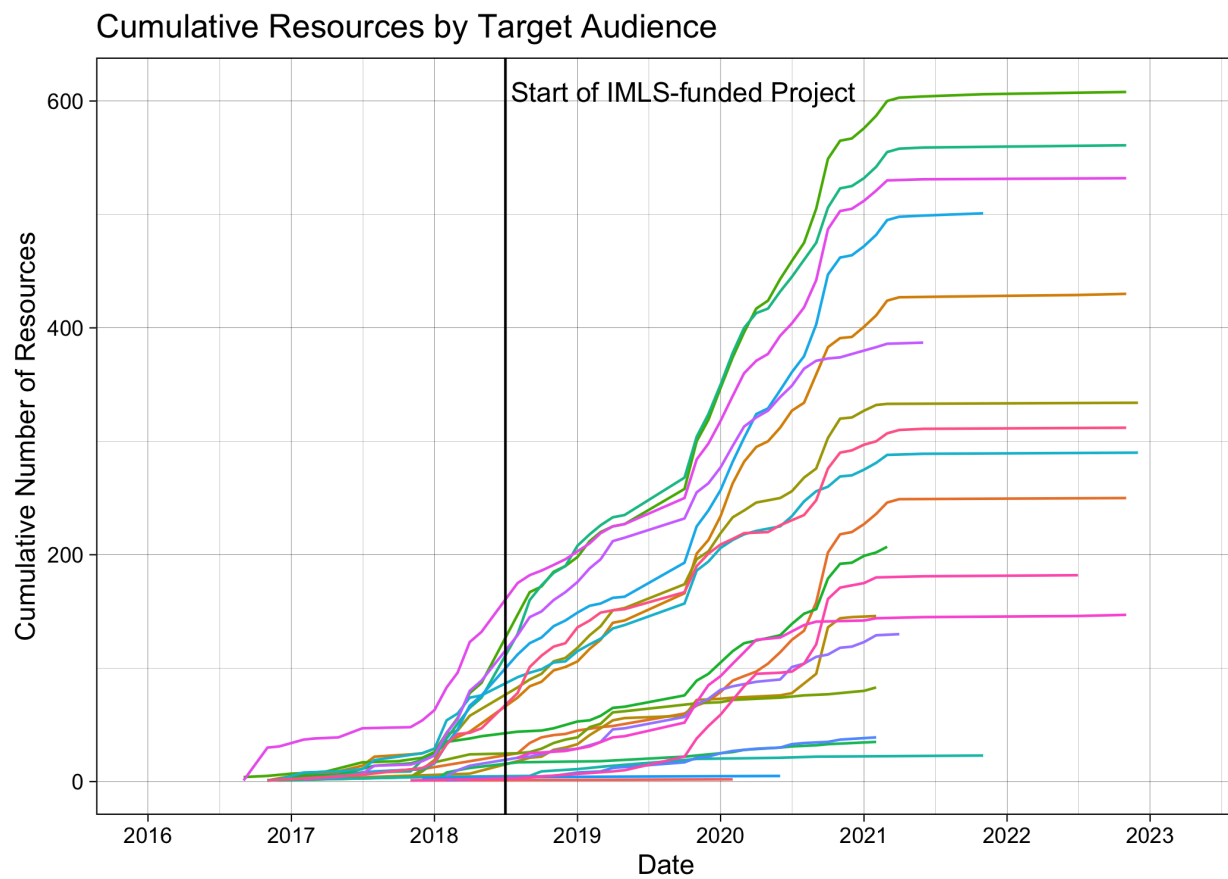
	Arts and Humanities
	character(0)
	Education
	Engineering
	Life Sciences
	Medicine and Health Sciences
	Physical Sciences and Mathematics
	Social and Behavioral Sciences

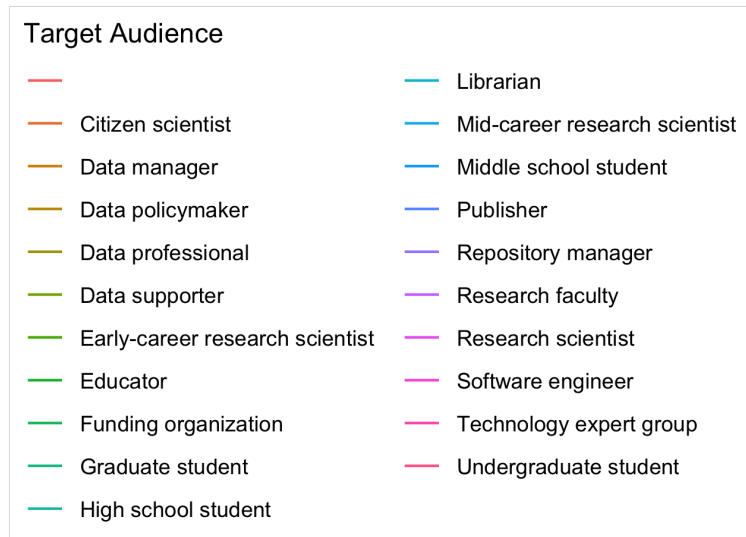
Target Audience Frequencies

target_audience	count
Early-career research scientist	608
Graduate student	561
Research scientist	532
Mid-career research scientist	501
Data manager	430
Research faculty	387
Data professional	334
Undergraduate student	312
Librarian	290
Citizen scientist	250
Educator	207
Technology expert group	182
Software engineer	147
Data policymaker	146
Repository manager	130
Data supporter	83
Publisher	39
Funding organization	35
High school student	23

(continued)

target_audience	count
Middle school student	5
	2





Keyword Frequencies

keywords	count
Data management	328
Data analysis	175
Data management planning	171
Data access	161
Data sharing	159
Data skills education	96
Data publication	95
Metadata	92
Data storage	90
Programming	89
Data visualization	83
Data archiving	80
Data reuse - Core Trustworthy Data Repositories Requirements	79
Data preservation	78
Data collection	75
Data management planning tools	74
Data formats	70
Satellite imagery	68

(continued)

keywords	count
Data lifecycle	67
Environmental management	65
Remote sensing	65
Open data	63
Data curation	57
Organizing data	55
Data citation	54
Data usage	54
Accessible data - FAIR Data Principle	52
Data coding	52
R software	52
Software management	52
Data visualization tools	51
Python	51
Re-usable data - FAIR Data Principle	50
Data handling	49
Data discovery and identification - Core Trustworthy Data Repositories Requirements	46
Community standards	45
Data access methods	43
Land management	39
Data policy	37
Metadata standards	37
Open access	37
Persistent Identifiers (PID)	36
Data documentation	35
Findable data - FAIR Data Principle	35
Licenses - Core Trustworthy Data Repositories Requirements	35
Data cleaning	33
Data modeling	33
Big data	30
Data stewardship	30
Data backup	29
Data transformation	29
Digital humanities	29
Interoperable data - FAIR Data Principle	29
Workflows - Core Trustworthy Data Repositories Requirements	29
Access rights	28
Agriculture data	28
Data quality	28
Geospatial data	28
Data ethics	27
	25
Capacity building	25
Copyright of data	25

(continued)

keywords	count
Data integration	25
FAIR metrics	25
Geographic Information System (GIS)	25
Ocean data	24
User guides for Earth science data	24
Cyberinfrastructure to enable FAIR data principles	23
Data quality - Core Trustworthy Data Repositories Requirements	23
Digital preservation workflow	23
Expert guidance - Core Trustworthy Data Repositories Requirements	23
Marine data	23
Social science data	23
Software preservation	23
Earth observation data	21
Flood mapping applications	21
Landcover applications	21
Workflow	21
Data governance	20
Data portals	20
Database design	20
Environmental change records	19
Humanities data	19
Scientific reproducibility	19
Structured Query Language (SQL)	19
Sustainable Development Goals (SDGs)	19
Business intelligence	18
Continuity of access - - Core Trustworthy Data Repositories Requirements	18
Discovery metadata	18
File naming	18
Research lifecycle	18
Conservation	17
Data management examples	17
Spreadsheets	17
ArcGIS	16
Climate data	16
Data wrangling	16
Database management	16
Open science	16
Research infrastructures	16
Biodiversity data	15
Data integrity and authenticity - Core Trustworthy Data Repositories Requirements	15
Electronic lab notebook (ELN)	15
Text and data mining	15

(continued)

keywords	count
Version control	15
Reusable software	14
Collaborative workflow	13
Data warehousing	13
Security - Core Trustworthy Data Repositories Requirements	13
Coastal data	12
Data management stories	12
Digital scholarship	12
Research project management tools	12
Synthetic aperture radar data (SAR)	12
Transformation tools for data	12
Data communication	11
Data literacy	11
Data privacy	11
Data research	11
Database models	11
Earthquake data	11
Historical data	11
Linked data	11
Network analysis	11
Preservation plan - Core Trustworthy Data Repositories Requirements	11
Scientific publications	11
Workshop curriculum	11
Cloud-based data	10
Confidentiality/Ethics - - Core Trustworthy Data Repositories Requirements	10
Data activities for Earth science data	10
How-to data recipes	10
Hydrologic data	10
Technical infrastructure - Core Trustworthy Data Repositories Requirements	10
Air quality data	9
Data conversion	9
Data use restrictions	9
Disaster applications	9
Drought monitoring	9
Elevation data	9
Fair use copyright doctrine	9
Food security data	9
Metadata registries	9
OpenRefine	9
Soils data assessment systems	9
Soils software and data management applications	9
Tabular data	9

(continued)

keywords	count
Clinical research data	8
Computer Science	8
Cultural heritage data	8
Forest mapping applications	8
Geoscience data	8
Programmatic digital preservation	8
Researcher engagement	8
Soil moisture analysis	8
Agriculture applications	7
Federal agency data management requirements (U.S.)	7
Metadata editors	7
Nutrition data	7
Open publications	7
Scientific record preservation	7
Shape files	7
Appraisal - Core Trustworthy Data Repositories Requirements	6
Archaeology data	6
Data loss	6
Freshwater habitat applications	6
Linux	6
PID data architecture	6
Research support skills	6
Scientific record archives	6
Animal experimental research data	5
Data management implementation services	5
Data user community	5
Digitization of data	5
Documented storage procedures - Core Trustworthy Data Repositories Requirements	5
Ecological processes	5
Genomics data	5
JavaScript	5
LiDAR technology	5
Linguistic data	5
Massive open online course (MOOC)	5
Mission/Scope - Core Trustworthy Data Repositories Requirements	5
Organizational infrastructure - - Core Trustworthy Data Repositories Requirements	5
Physical samples (Earth Science)	5
Provenance	5
Research Data Services (RDS)	5
Research transparency	5
Statistical modeling	5
Unix	5

(continued)

keywords	count
Biological data	4
Scholarly publishing	4
Statistical analysis	4
Water quality	4
Aichi Biodiversity Targets (ABTs)	3
Cyberinfrastructure for science	3
Gridded data	3
Reference management	3
Social anthropology data	3
Sustainable cities	3
Swath bathymetry data	3
Anonymisation of data	2
Building information modeling (BIM)	2
Climate resilience	2
Context - Core Trustworthy Data Repositories Requirements	2
Crop monitoring	2
Cyberinfrastructure and the research library	2
Data linkage	2
Data transfer	2
Discoverable software	2
Experimental design	2
Format conversion	2
Informed consent	2
International Geo Sample Number (IGSN)	2
Library science	2
Machine learning	2
Machine-actionable data management plans (DMPs)	2
Natural language processing	2
Responsible research and innovation (RRI)	2
Revit file format	2
Software sharing	2
Taxonomic data	2
3 D modeling	1
Biogeochemical data	1
Climate risk assessment applications	1
Community climate system model (CCSM)	1
Creative arts data	1
Data Seal of Approval (DSA)	1
Data architecture	1
Data containers	1
Data sustainability	1
Electronic theses and dissertations (ETD)	1
Encryption	1
European diversity	1

(continued)

keywords	count
Fossil data	1
Geo for All	1
Graphical User Interface (GUI)	1
Grey literature	1
Groundwater monitoring	1
Hazard identification and risk analysis (HIRA)	1
High performance computing (HPC)	1
In-vivo experimental data	1
Indexing	1
Legal research	1
Multidimensional data (Scientific)	1
Numeric data	1
Open business models	1
Open principles for geospatial education	1
Organismal data	1
Organizational change	1
Paleontological data	1
Photogrammetry	1
Probabilistic data linkage	1
Qualitative data analysis	1
Relational database	1
Restricted data	1
Risk analysis	1
River basin management	1
Science gateways	1
Scientific Writing	1
Self-describing data formats	1
Software citation	1
Spatial Capture-Recapture modeling	1
Specimen data	1
Survey data	1
System administration	1
Technology hotel	1
Tidy data	1

