

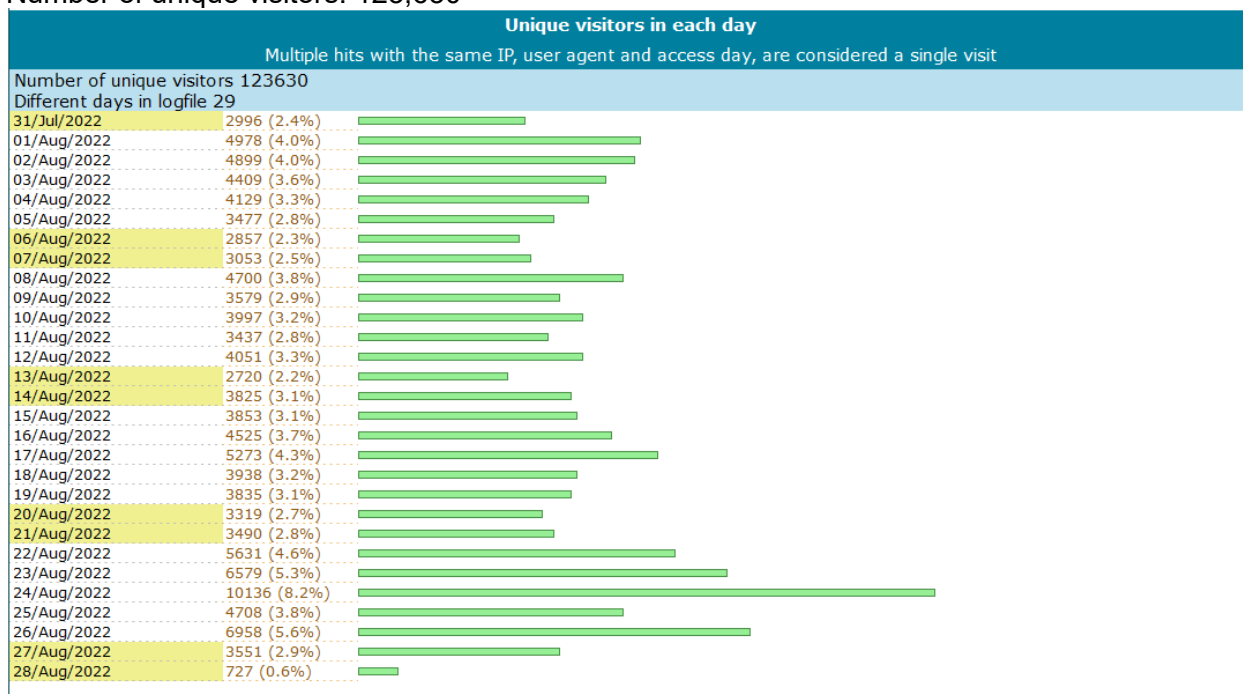
Access and Use Statistics and FAIR metrics on Metabolomics Workbench (MW)

Approach: We used both custom analysis of apache logs and Visitors tool (<http://www.hping.org/visitors/>) on combined access_log and ssl_access_log. vsftpd log was analyzed separately.

Access and use statistics are based on analysis of the log over about one month (Aug 01 – Aug 28, 2022, unless indicated otherwise).

Overall access to the MW website

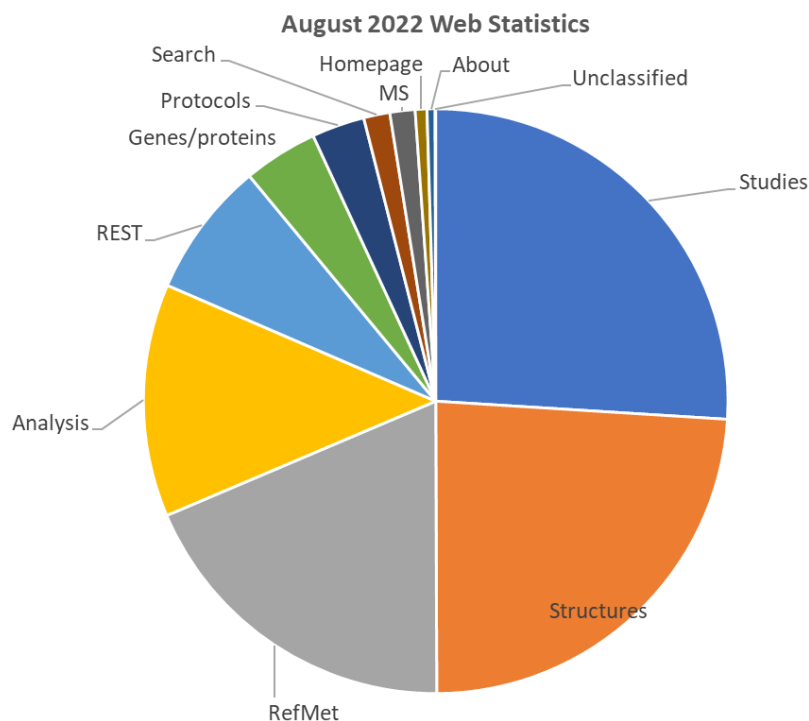
Number of unique visitors: 123,630



Resource types (Major categories)

Category	August 2022 Web Statistics (31 days)	Hits per day
Studies	1097761	35412
Structures	1011719	32636
RefMet	788463	25434
Analysis	543812	17542
REST	317354	10237
Genes/proteins	173855	5608
Protocols	123529	3985
Search	61099	1971
MS	58800	1897
Homepage	27535	888

About	18764	605
Unclassified	1241	40



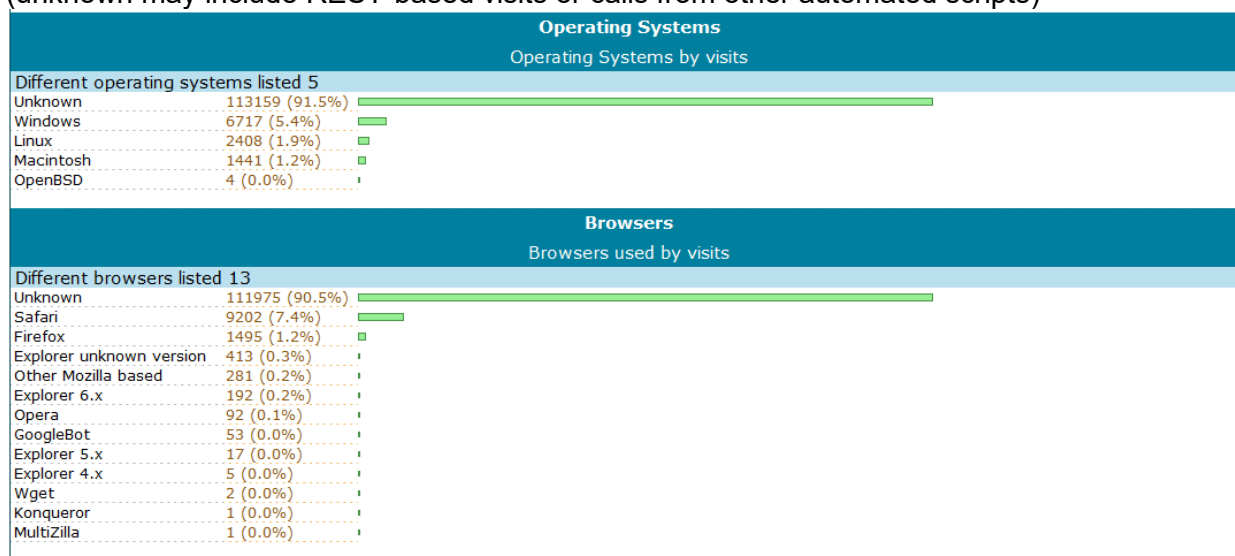
Top page hits over Aug 1-28, 2022:

Type of page	End-point	Count
Studies	/data/DRCCMetadata.php	790608
REFMET	/databases/refmet/name_to_refmet_new_min.php	523774
Statistical analysis	/data/metabolite_id_show.php	361963
Statistical analysis	/data/show_metabolite_data_by_factors.php	261543
Structure data	/data/StructureData.php	185007
REFMET	/rest/refmet	180821
Statistical analysis	/data/ttest1.php	113645
Proteome	/databases/proteome/MGP_table.php	94396
Statistical analysis	/data/ttest1_grid.php	86554
Studies	/rest/study	77579
REFMET	/databases/refmet/refmet.php	64928
Statistical analysis	/data/zscore_calc.php	63537
Statistical analysis	/data/anova1.php	62997
Pathways	/data/show_h_pathway_metabolites_reactome.php	56701
Metabolite-Gene	/databases/proteome/MGP_detail.php	54744
Protocols	/protocols/protocoldetails.php	54221
Massbank	/data/massbank.php	50033
Search	/search/sitesearch.php	47282
Protocols	/protocols/general.php	40948

REFMET	/databases/refmet/abbrev_generic.php	39366
Study summary	/data/DRCCStudySummary.php	36577
Search	/data/s3.php	36562
Files	/data/file_extract_7z.php	32910
REST	/rest/protein	32780
REFMET	/databases/refmet/refmet_details.php	29519

OS and browsers

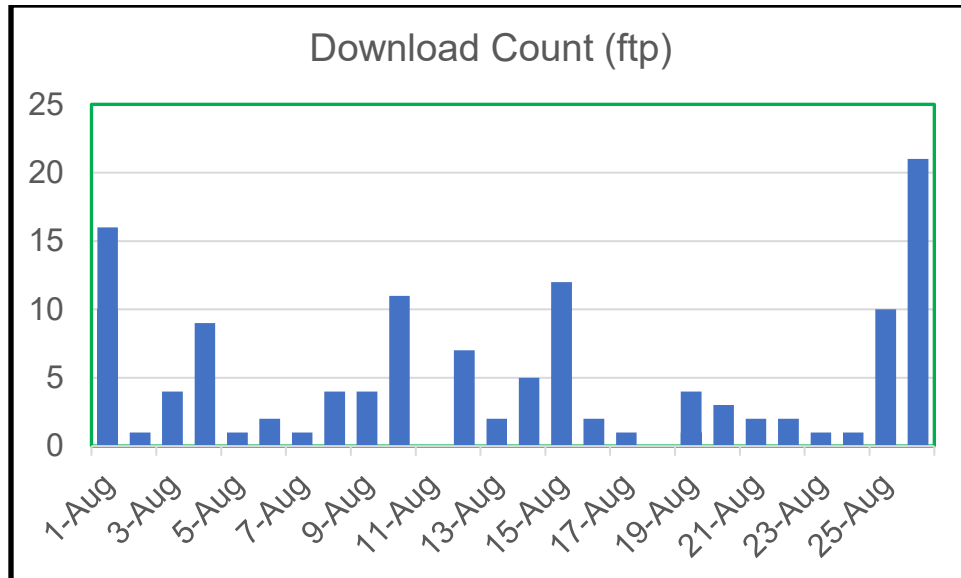
(unknown may include REST-based visits or calls from other automated scripts)



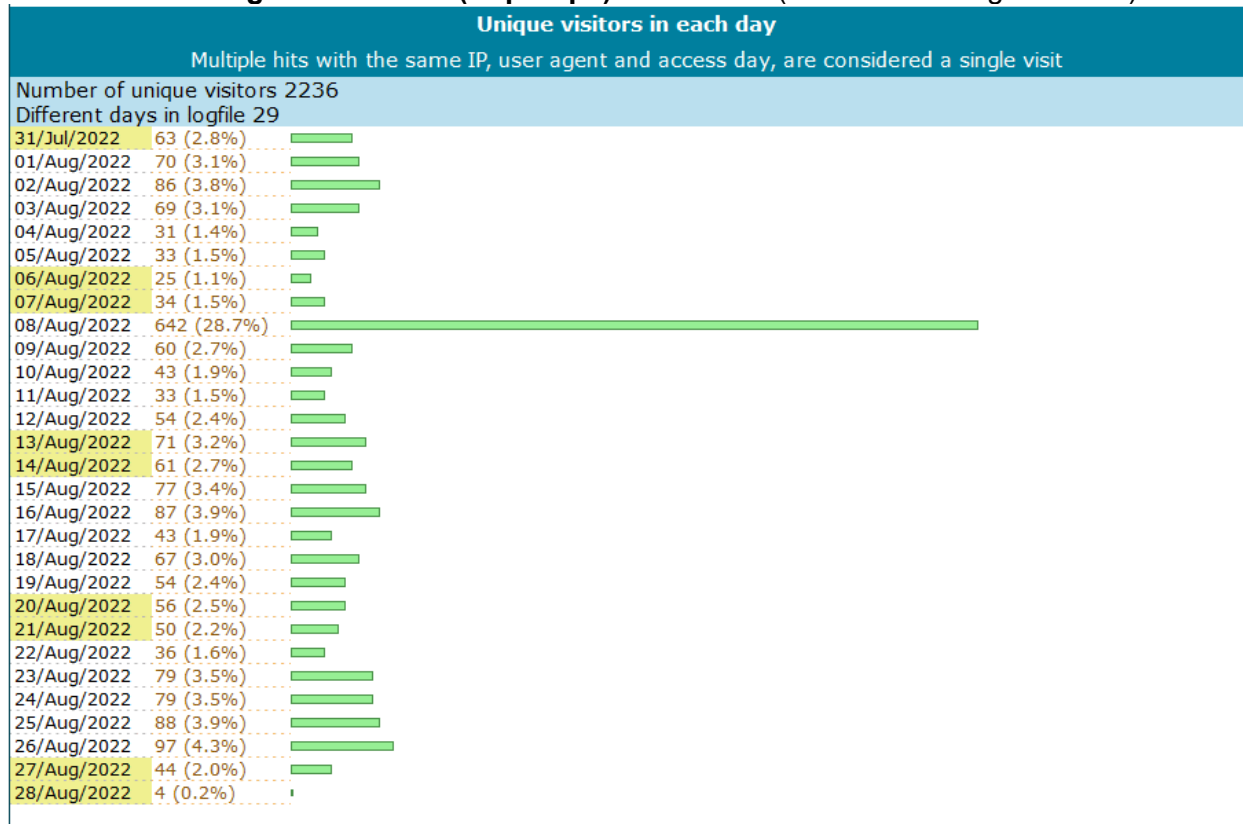
Download statistics

Amount of data downloaded: about 14 TB

FTP-based downloads: Total 127



Downloads through the website (http/https): Total 2236 (Generated using "Visitors")

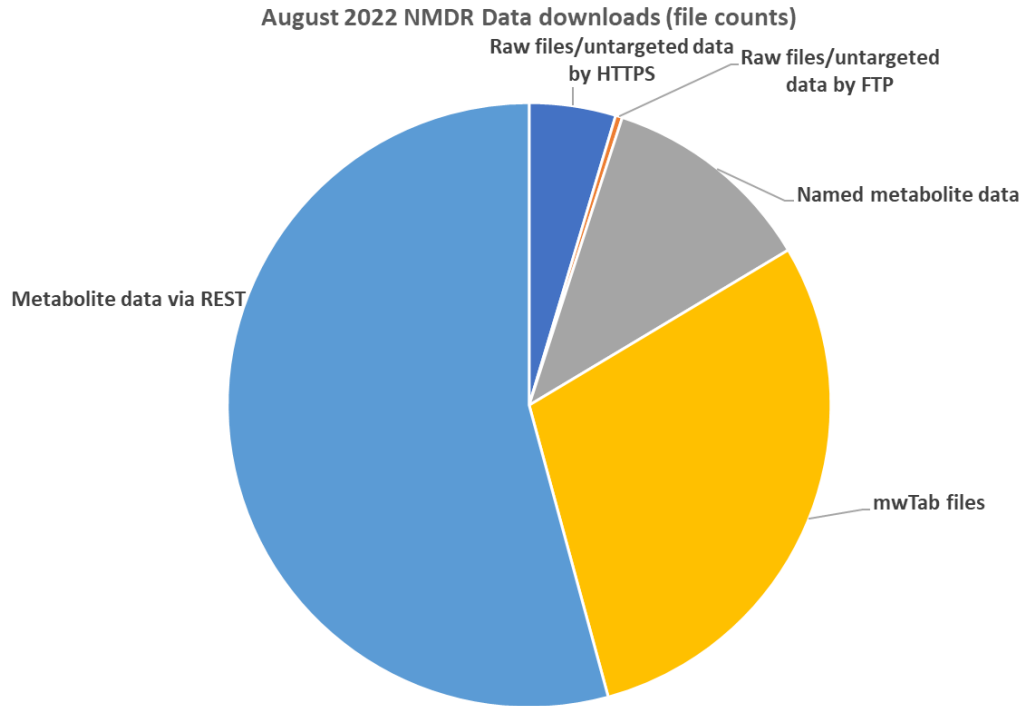


Top downloads requested:

Requested pages		
Page requests ordered by hits		
Different pages requested 1241		
1)	74647	/studydownload/ST001757_serum_rawdata.zip
2)	74436	/studydownload/ST001758_brain_rawdata.zip
3)	66026	/studydownload/ST001756_feces_rawdata.zip
4)	3049	/studydownload/ST001873_rawdata.zip
5)	175	/studydownload/ST000923_AN001513_Results.txt
6)	157	/studydownload/ST000758_AN001190_Results.txt
7)	154	/studydownload/ST001192_AN001984_Results.txt
8)	90	/studydownload/ST000775_AN001222_Results.txt
9)	78	/studydownload/ST000879_AN001429_Results.txt
10)	54	/studydownload/ST000746_AN001169_Results.txt
11)	51	/studydownload/ST000622_AN000954_Results.txt
12)	50	/studydownload/ST000919_AN001506_Results.txt
13)	50	/studydownload/ST001272.zip
14)	43	/studydownload/ST000879_AN001432_Results.txt
15)	27	/studydownload/ST001000_AN001879_Results.txt
16)	27	/studydownload/ST002120_Lung.zip
17)	25	/studydownload/ST000867_Results_combined_GCMS_LCMS_Data.xlsx
18)	23	/studydownload/ST001000.zip
19)	14	/studydownload/ST002087_Data.zip
20)	14	/studydownload/ST000578_AN000888_Results.txt

Mode of download (based on 31 days)

Type	August 2022 NMDR Data downloads (file counts)
Raw files/untargeted data by HTTPS	4007
Raw files/untargeted data by FTP	315
Named metabolite data	9891
mwTab files	25387
Metabolite data via REST	46905



Access statistics of our GitHub page for Jupyter notebooks repository (covers 6/1/2022 – 09/06/2022)

<https://github.com/metabolomicsworkbench/jupyter-notebooks>

Git clones: Unique: 44 Total: 44

Accesses by path:

path	count	uniques
/metabolomicsworkbench/jupyter-notebooks Total	62	32
/metabolomicsworkbench/jupyter-notebooks/blob/master/MWPerformClusteredHeatMapAnalysis.ipynb Total	1	1
/metabolomicsworkbench/jupyter-notebooks/blob/master/MWPerformDataNormalization.ipynb Total	13	13
/metabolomicsworkbench/jupyter-notebooks/blob/master/MWPerformLinearDiscriminantAnalysis.ipynb Total	7	5
/metabolomicsworkbench/jupyter-notebooks/blob/master/MWPerformPartialLeastSquaresDiscriminantAnalysis.ipynb Total	1	1
/metabolomicsworkbench/jupyter-notebooks/blob/master/MWPerformPrincipalComponentAnalysis.ipynb Total	3	3
/metabolomicsworkbench/jupyter-notebooks/blob/master/MWPerformRandomForestAnalysis.ipynb Total	8	6
/metabolomicsworkbench/jupyter-notebooks/blob/master/MWPerformRelativeLogAbundanceAnalysis.ipynb Total	9	7
/metabolomicsworkbench/jupyter-notebooks/blob/master/MWPerformVolcanoPlotAnalysis.ipynb Total	23	21
/metabolomicsworkbench/jupyter-notebooks/blob/master/MWRestAPICompoundDataExample.ipynb Total	9	9
/metabolomicsworkbench/jupyter-notebooks/blob/master/MWRestAPIGeneDataExample.ipynb Total	3	3
/metabolomicsworkbench/jupyter-notebooks/blob/master/MWRestAPINamedMetabolitesResultsExample.ipynb Total	2	2
/metabolomicsworkbench/jupyter-notebooks/blob/master/MWRestAPIProteinDataExample.ipynb Total	7	7
/metabolomicsworkbench/jupyter-notebooks/blob/master/MWRestAPIStudyDataExample.ipynb Total	6	2
/metabolomicsworkbench/jupyter-notebooks/blob/master/MWSampleData1.csv Total	2	2
/metabolomicsworkbench/jupyter-notebooks/blob/master/README.md Total	1	1

referrer	count
github.com Total	15
Google Total	63
metabolomicsworkbench.org	
Total	94
notebooks.gesis.org Total	2

<https://github.com/metabolomicsworkbench/MetENP>

Git clones: Unique: 70 Total: 86

path	count	uniques
/metabolomicsworkbench/MetENP Total	245	85
/metabolomicsworkbench/MetENP/blob/main/binder/environment.yml Total	8	4
/metabolomicsworkbench/MetENP/blob/main/binder/postBuild Total	5	4
/metabolomicsworkbench/MetENP/blob/main/data/ls_path.RData Total	2	2
/metabolomicsworkbench/MetENP/blob/main/DESCRIPTION Total	6	5
/metabolomicsworkbench/MetENP/blob/main/environment.yml Total	4	3
/metabolomicsworkbench/MetENP/blob/main/LICENSE Total	1	1
/metabolomicsworkbench/MetENP/blob/main/man/enzyme_gene_info.Rd Total	2	1
/metabolomicsworkbench/MetENP/blob/main/NAMESPACE Total	1	1
/metabolomicsworkbench/MetENP/blob/main/R/anova_ana.R Total	1	1
/metabolomicsworkbench/MetENP/blob/main/R/compoundinfo.R Total	2	2
/metabolomicsworkbench/MetENP/blob/main/R/convert_refmet.R Total	6	4
/metabolomicsworkbench/MetENP/blob/main/R/metclassenrichment.R Total	12	10
/metabolomicsworkbench/MetENP/blob/main/README.md Total	9	3
/metabolomicsworkbench/MetENP/blob/main/vignettes/Case%20study%201_%20Study%20ST000915.html Total	8	4
/metabolomicsworkbench/MetENP/blob/main/vignettes/MeTENP_ST000915.ipynb Total	2	1
/metabolomicsworkbench/MetENP/blob/main/vignettes/MetENP_vignette_Jupyter_notebook.ipynb Total	7	3
/metabolomicsworkbench/MetENP/blob/main/vignettes/MetENP_vignette.rmd Total	9	5

referrer	count
bdcw.org Total	4
biorexiv.org Total	68
github.com Total	61
Google Total	136
metabolomicsworkbench.org Total	27

<https://github.com/metabolomicsworkbench/MetENPAppyter>

Git clones: Unique: 5 Total: 6

path	count	uniques
/metabolomicsworkbench/MetENPAppyter Total	5	4
/metabolomicsworkbench/MetENPAppyter/blob/main/R/MetENP/R/MetENP Total	2	2
/metabolomicsworkbench/MetENPAppyter/tree/main/R/MetENP Total	2	2
/metabolomicsworkbench/MetENPAppyter/tree/main/R/MetENP/R Total	2	2

referrer	count
github.com Total	2
Google Total	9

The number of accesses to the Jupyter/Binder entry page at:




<https://www.metabolomicsworkbench.org/data/AnalyzeUsingJupyterNotebooks.php>

The number of hits to the entry page is low- about 20 per week. This doesn't necessarily mean that users are clicking on the external Binder/Github links on this page, though.

FAIRShake metrics

Precalculated stats on MW: <https://fairshake.cloud/project/85/assessments/>


Project Assessments (6697)

Assessment			Metrics								
Target	Rubric		Globally unique identifier	Persistent identifier	Machine-readable metadata	Standardized metadata	Resource identifier in metadata	Resource discovery through web search	Open, Free, Standardized Access protocol	Protocol to access restricted content	Persistence of resource and metadata
Fatb Induction Experiment (FatBIE)	FAIR metrics by fairmetrics.org		no (0.00)	no (0.00)	yes (1.00)	yes (1.00)	yes (1.00)	yes (1.00)	yes (1.00)		no (0.00)
Intestinal Samples II pre/post transplantation	FAIR metrics by fairmetrics.org		no (0.00)	no (0.00)	yes (1.00)	yes (1.00)	yes (1.00)	yes (1.00)	yes (1.00)		no (0.00)
Metabolomic analysis of mouse embryonic fibroblasts, embryonic stem cells, and induced pluripotent stem cells	FAIR metrics by fairmetrics.org		no (0.00)	no (0.00)	yes (1.00)	yes (1.00)	yes (1.00)	yes (1.00)	yes (1.00)		no (0.00)

Apparently, the FAIRshake tool that generated the above statistics is not fully mature and requires manual curation. MW uses document object identifiers (DOI) for the projects, which is included (referred to as Persistent Identifier) in the metadata submitted to the CFDE portal. Through the DOI, one can access all the publicly available data on the MW website.

https://fairshake.cloud/digital_object/6578/assessments/

Tags: NIHcommonfund

Assessment			Metrics								
Rubric	Project		The repository provides contact information for staff to enable users with questions or suggestions to interact with repository experts.	Tools that can be used to analyze each dataset are listed on the corresponding dataset pages.	The repository maintains licenses to manage data access and use.	The repository hosts data and metadata according to a set of defined criteria to ensure that the resources provided are consistent with the intent of the repository.	The repository provides documentation for each resource to permit its complete and accurate citation.	A description of the methods used to acquire the data is provided.	Version information is provided for each resource, where available.	The structure of the repository permits efficient discovery of data and metadata by end users.	The repository uses a standardized protocol to permit access by users.
The FAIRshake repository rubric	NIH Data Sharing Repositories		yes (1.00)	yes (1.00)	no (0.00)	yes (1.00)	yes (1.00)	yes (1.00)	yes (1.00)	yes (1.00)	yes (1.00)

The above tables lists “no” for license, but we have stated the terms of use. In the future, we will decide the license term (CC BY-NC, CC BY-NC-SA or something allowing commercialization of derivative work) and update the “Terms of use” page.



METABOLOMICS WORKBENCH

[Log in / Register](#)



[Home](#) | [Data Repository](#) | [Databases](#) | [Protocols](#) | [Tools](#) | [Training / Events](#) | [About](#) | [Search](#)

[Overview](#) | [Working Groups](#)  | [Publishing Partners](#) | [How to Cite](#) | [How to Link](#) | [Terms of Use](#) | [Personnel](#) | [Contact](#) | [Sitemap](#)

Terms of Use

The Metabolomics Workbench is provided by the NIH Common Fund's National Metabolomics Data Repository (NMDR) on an "as is" basis, without warranty or representation of any kind, express or implied. The content of the Metabolomics Workbench is protected by international copyright, trademark and other laws. You may download articles and web pages from this site for your personal, non-commercial use only, provided that you keep intact all authorship, copyright and other proprietary notices. If you use the Metabolomics Workbench, you accept these terms. The NMDR reserves the right to modify these terms at any time.