SCCWRP #1276

frontiers | Frontiers in Environmental Science

ORIGINAL RESEARCH
published: 30 June 2022
doi: 10.3389/fenvs.2022.912107

Current State of Microplastic Pollution Research Data: Trends in Availability and Sources of Open Data

Tia Jenkins[1], Bhaleka D. Persaud[1], Win Cowger[2], Kathy Szigeti[3], Dominique G. Roche[4,5], Erin Clary[6], Stephanie Slowinski[1], Benjamin Lei[1], Amila Abeynayaka[7], Ebenezer S. Nyadjro[8,9], Thomas Maes[10], Leah Thornton Hampton[11], Melanie Bergmann[12], Julian Aherne[13], Sherri A. Mason[14], John F. Honek[15], Fereidoun Rezanezhad[1], Amy L. Lusher[16], Andy M. Booth[17], Rodney D. L. Smith[15] and Philippe Van Cappellen[1]*

[1]Department of Earth and Environmental Sciences, University of Waterloo, Waterloo, ON, Canada, [2]Moore Institute for Plastic Pollution Research, Long Beach, CA, United States, [3]Davis Centre Library, University of Waterloo, Waterloo, ON, Canada, [4]Department of Biology, Carleton University, Ottawa, ON, Canada, [5]Institute of Biology, University of Neuchâtel, Neuchâtel, Switzerland, [6]Digital Research Alliance of Canada, Ottawa, ON, Canada, [7]Institute for Global Environment Strategies (IGES), Kanagawa, Japan, [8]National Oceanic and Atmospheric Administration (NOAA) National Centers for Environmental Information (NCEI), Stennis Space Center, Starkville, MS, United States, [9]Northern Gulf Institute, Mississippi State University, Stennis Space Center, Starkville, MS, United States, [10]GRID-Arendal, Arendal, Norway, [11]Southern California Coastal Water Research Project (SCCWRP), Costa Mesa, CA, United States, [12]Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung, Bremerhaven, Germany, [13]School of Environment, Trent University, Peterborough, ON, Canada, [14]The Behrend College, Pennsylvania State University, Erie, PA, United States, [15]Department of Chemistry, University of Waterloo, Waterloo, ON, Canada, [16]Norwegian Institute for Water Research, Oslo, Norway, [17]SINTEF Ocean, Trondheim, Norway

OPEN ACCESS

Edited by:
Lingzhan Miao,
Hohai University, China

Reviewed by:
Hong Li,
Chongqing University, China
Kathryn Hassell,
RMIT University, Australia
Longfei Wang,
Hohai University, China

****Correspondence:***
Tia Jenkins
t4jenkins@uwaterloo.ca
orcid.org/0000-0003-2449-6795

Specialty section:
This article was submitted to Freshwater Science, a section of the journal Frontiers in Environmental Science

Received: *04 April 2022*
Accepted: *23 May 2022*
Published: *30 June 2022*

Citation:
Jenkins T, Persaud BD, Cowger W, Szigeti K, Roche DG, Clary E, Slowinski S, Lei B, Abeynayaka A, Nyadjro ES, Maes T, Thornton Hampton L, Bergmann M, Aherne J, Mason SA, Honek JF, Rezanezhad F, Lusher AL, Booth AM, Smith RDL and Van Cappellen P (2022) Current State of Microplastic Pollution Research Data: Trends in Availability and Sources of Open Data. Front. Environ. Sci. 10:912107. doi: 10.3389/fenvs.2022.912107

The rapid growth in microplastic pollution research is influencing funding priorities, environmental policy, and public perceptions of risks to water quality and environmental and human health. Ensuring that environmental microplastics research data are findable, accessible, interoperable, and reusable (FAIR) is essential to inform policy and mitigation strategies. We present a bibliographic analysis of data sharing practices in the environmental microplastics research community, highlighting the state of openness of microplastics data. A stratified (by year) random subset of 785 of 6,608 microplastics articles indexed in Web of Science indicates that, since 2006, less than a third (28.5%) contained a data sharing statement. These statements further show that most often, the data were provided in the articles' supplementary material (38.8%) and only 13.8% via a data repository. Of the 279 microplastics datasets found in online data repositories, 20.4% presented only metadata with access to the data requiring additional approval. Although increasing, the rate of microplastic data sharing still lags behind that of publication of peer-reviewed articles on environmental microplastics. About a quarter of the repository data originated from North America (12.8%) and Europe (13.4%). Marine and estuarine environments are the most frequently sampled systems (26.2%); sediments (18.8%) and water (15.3%) are the predominant media. Of the available datasets accessible, 15.4% and 18.2% do not have adequate metadata to determine the sampling location and media type, respectively. We discuss five recommendations to strengthen data sharing practices in the environmental microplastic research community.

Keywords: microplastics, bibliometric analysis, data repository, data availability statement, data management, data sharing, environmental research, plastic

INTRODUCTION

There is an increasing awareness of microplastics in the environment and their potential negative consequences for water security, biodiversity, ecosystem services, human health and well-being (Bergmann et al., 2015; Barboza et al., 2018; Li et al., 2018; Provencher et al., 2020; Woods et al., 2021; Stokstad, 2022). Along with other novel entities, microplastic pollution is now considered to exceed safe planetary boundaries (Persson et al., 2022). This awareness has spurred a surge in research on microplastics, including their occurrence and environmental distributions, chemical and physical properties, fate and transport (Domercq et al., 2022), impacts on biota and ecosystems (Abeynayaka and Norihiro, 2019; Covernton et al., 2019; Jacques and Prosser, 2021; Tekman et al., 2022) and integration into life cycle inventories and impact assessment (Abeynayaka and Norihiro, 2019; Woods et al., 2021). The increasing interest in microplastics is reflected in the number of published peer-reviewed articles and news articles (Ryan, 2015; Cowger et al., 2020; Can-Güven, 2021). The rapid growth of publications on microplastic pollution since the turn of the century is primarily associated with research on marine environments, freshwater bodies, wastewater, and fate and transport of microplastics, with publications spanning eighty-seven countries across the globe (Can-Güven, 2021). Simultaneously, funding to support microplastics research has increased in the past decade (Maes et al., 2019). For instance, the Government of Canada has made the detection and characterization of microplastics a priority area for research funding to develop the knowledge base and research capacity required to support Canada's Plastics Science Agenda (CaPSA) (Environment and Climate Change Canada, 2019; NSERC, 2020). In the United States, the National Oceanic and Atmospheric Administration (NOAA) Marine Debris and NOAA Sea Grant programmes offer research funding that focuses on plastic pollution. Many state sea grant programs also now include plastic pollution as a priority area (Sea Grant, 2018; NOAA and NECEI, 2022).

Researchers are developing new approaches to isolate, count, and measure microplastics in different environmental settings to characterise the global distribution of microplastic (*e.g.,* see Arctic Monitoring and Assessment Programme (AMAP) report; Wayne State's Smart Management of Microplastic Pollution), which is critical to guide the state of our knowledge on sources, fate, and effect of microplastics, and to facilitate and assess effective policy decision-making. For example, the United Nations passed a major global resolution on plastic pollution in March 2022 (Stokstad, 2022), while the State of California adopted a state-wide microplastics strategy in February 2022 (State of California, 2022). To ensure good decision-making and to enforce these policies, microplastic data must be made FAIR (Findable, Accessible, Interoperable, and Reusable), as they are key to the process. To advance research, protect funder investments in data collection, enable policy development, and support public interest into the human and environmental health impacts of microplastics (Koelmans et al., 2019; Cowger et al., 2020; Igalavithana et al., 2022), research data must be properly curated, deposited and preserved in adherence with the FAIR guiding principles (Wilkinson et al., 2016).

As more and more data on microplastics are acquired and as policies begin to emerge around the world (e.g. Stokstad, 2022), it is important that scientists are able to conduct meta-analyses, confirm reproducibility, and meaningfully compare data from different studies (Cowger et al., 2020; Provencher et al., 2020; Brandes et al., 2021). The international workshop on microplastic particles organised by the Joint Group of Experts on the Scientific Aspects of Marine Environmental Protection (GESAMP) articulated the fundamental gaps in microplastics standardisation must be filled in order to enable the comparison and merging of data from researchers from across various geographic regions (GESAMP, 2019). These fundamental gaps include data capture standards, quality control practices, data storage and sharing, as well as reporting and dissemination. A decade later, microplastics datasets are generated rapidly and stored in a variety of formats, from open source to proprietary, and data range in size, from kilobytes to terabytes and many of these datasets are still not finable (Brandes et al., 2021). Nonetheless, subsequent international activities have been initiated to address some of these gaps, including coordination of global and regional efforts to characterize plastics pollution by generating guidelines for sampling and reporting that will minimize the duplication of work (e.g., Global Partnership on Marine Litter, Japan's Ministry of Environment, OSPAR Commission, NOAA's NCEI Microplastics). In addition, recently a number of microplastics-focused data repositories have also been created in an effort to homogenise subsets of data (Morgan Stanley, National Geographic, University of Georgia, and National Oceanic Atmospheric Administration, 2010; Tekman et al., 2020; NOAA NCEI Microplastics n. d; EU EMODNET, 2017), but how much and what types of microplastics data are generated by, and readily available for the academic research community remains unclear. An assessment of the state of microplastics research data accessibility would assist researchers and stakeholders in identifying best data management practices in the field.

In this study, we explore the extent to which the data that underpin environmental microplastics research articles are openly shared. Two strategies were adopted to identify and locate open datasets: 1) we reviewed the data sharing statements in a representative subset of peer-reviewed publications on environmental microplastics, and 2) we undertook a comprehensive search of relevant online data repositories. Based on our findings, we highlighted five practices that researchers in the microplastics community can readily implement to advance data sharing in this emerging field.

DATA SOURCES AND METHODOLOGY

Analysis 1: Publications on Microplastics With Open Data

The methods outlined by Read et al. (2021) and Roche et al. (2022b) were adapted to determine if authors of microplastic research articles shared the underlying data. A Web of Science database search was performed using the Boolean phrase (microplastic OR microplastics)

TABLE 1 | The selection criteria and number of peer-reviewed journal articles that were identified and used for data analysis in this study.

Selection Criteria	Number of articles
Web of Science search (microplastic OR microplastics) between 1964 and 2021	6608
Stratified random sampling of up to 100 articles per year between 1964 and 2021	1045
Abstract only—removed from sample set	2
Articles not related to environmental microplastics—removed from sample set	256
Articles that could not be accessed—removed from sample set	1
Articles that were retracted—removed from sample set	1
Total articles removed from sample set between 1964 and 2021	260
Percentage of articles removed from sample set	24.9%
Number of articles included in the final sample set between 2006 and 2021	785

together with the "All Fields" option. Only English-language articles published between 1964 and 2021 were included, which yielded 6,608 articles. A stratified random sample selection of these articles was conducted in *R* (version 4.0.3) using the dplyr package (version 1.0.2) to select 100 studies per year. For years in which fewer than 100 studies were published, all articles were considered. The resulting subset consisted of 1,045 articles (15.8%). A number of articles were removed from this subset after manual inspection (n = 260; 24.9%) as these articles dealt with unrelated topics (for example "microplastic deformation" of metal alloys), or with topics not directly dealing with environmental samples (such as microbial colonisation of microplastics), or they were perspective-style or review papers. Additionally, articles that were inaccessible, retracted, or consisted solely of an abstract were removed. After the manual inspection and removal of articles that were not environmental microplastics related, a total of 785 peer-reviewed publications (11.9%) were included in the final assessment (**Table 1**).

Each of the selected articles was examined to determine 1) whether a data sharing statement was included in the article and, if so, 2) what the nature of the data sharing statement was. The nature of data sharing statements was categorised as: (i) available upon request via the author(s), (ii) available in a data repository, (iii) available in the supplementary files, (iv) no data were used, (v) data will be made available at a future date, (vi) no evidence of data sharing, (vii) data are considered sensitive, or (viii) data are available in the article.

We ensured that our metrics for percent of studies with data statements would be reproducible using simulations to determine the number of studies we needed to assess. A thousand simulations of subsampling from a two-class set (article does/does not have a data statement) with uniform probability distributions were measured by calculating the high mean absolute errors of the class percentages at the 95% quantile of the simulation distribution. Imposing the minimum number of studies to review at 100 per year yielded a maximum mean absolute error in class size of ±9%. Thus, if data accessibility changed by at least 18% over our period of study or in future studies we would be highly likely to identify the temporal change.

Analysis 2: Microplastics Datasets in Data Repositories

To assess the availability of data, the google dataset, DataONE Data One (2015) and OpenAIRE (2013) discovery portals were searched in October 2021 using the same search terms as used in Analysis 1 (i.e., microplastic OR microplastics) to identify available microplastics datasets. This search generated 10 repositories (**Table 2**). This was followed by a site-specific search of the 10 repositories to assess data access and metadata. The search engine of each repository was queried using the term "microplastic*" or the Boolean phrase "microplastic OR microplastics" when the search interface did not support the use of a wildcard. The search was not restricted to any specific time period and duplicated datasets ($n = 21$) were removed from the final sample of datasets. For each dataset, we recorded the following attributes: repository, year of publication, DOI, study site, environmental media type, keywords, and whether the dataset was linked to a journal article (yes/no). For each repository, we noted the disciplinary data it accepts (**Table 2**) and whether the repository was CoreTrustSeal, 2022 (CTS) certified as a trustworthy data repository as of February 2022 (**Table 2**), according to the CTS Certified Repository website. All metadata were recorded in a Microsoft Excel spreadsheet and OriginPro 2020 software was used to visualise the findings.

RESULTS

Microplastics Data Sharing Trends in Peer-Reviewed Articles

Of the final 785 articles analysed, 224 (28.5%) contained data sharing statements in the body of the article (**Figure 1A**). Prior to 2013, only eight out of 31 articles included data sharing statements. Since then, the numbers have steadily increased, with approximately half of all articles published after 2019 containing a data statement (**Figure 1B**). However, the proportion of articles with data sharing statements did not increase further between 2019 and 2021.

Further evaluation of the 224 data sharing statements (**Figure 2**) showed that authors most frequently shared the data associated with their article in the form of supplementary materials (n = 87; 38.8%) or stated that they had included all their data in the main body of the article (n = 60; 26.8%). In a small number of cases, the data underlying the publication were classified as sensitive and could not be shared (n = 2; 0.9%). A similar small number of statements indicated that the data would be made available in the future (n = 2; 0.9%). Others referred the reader to the corresponding author to

TABLE 2 | Selected research data repositories used in this study.

Repository name	Acronym	Discipline(s)	CTS Certified	No. of Datasets (*n*)
Dryad	Dryad	General, ecology and evolutionary biology	Not listed	21
Environmental Data Initiative Portal	EDI Portal	Environmental and ecological data	Not listed	14
Environmental Information Data Centre	EIDC	Datasets related to terrestrial and freshwater sciences	16 August 2019	4
Figshare	Figshare	General	Not listed	29
Harvard Dataverse	Harvard DV	General, social sciences	Not listed	5
Mendeley Data	Mendeley	General	Yes (expired)	75
Pangaea	Pangaea	Earth system science	17 June 2019	90
Polar Data Catalogue	PDC	Focus on cold and high latitude regions	16 February 2021	15
SEA scieNtific Open data Edition	SEANOE	Marine sciences	Not listed	4
Zenodo	Zenodo	General	Not listed	22

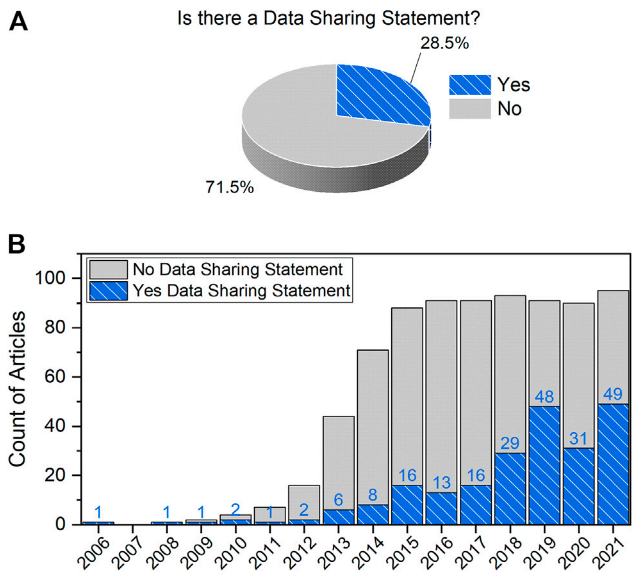

FIGURE 1 | Proportion of data sharing statements **(A)** overall and **(B)** annually between 2006 and 2021 in the peer-reviewed publications. Numbers displayed on the bars in panel b refer to the yearly number of data sharing statements.

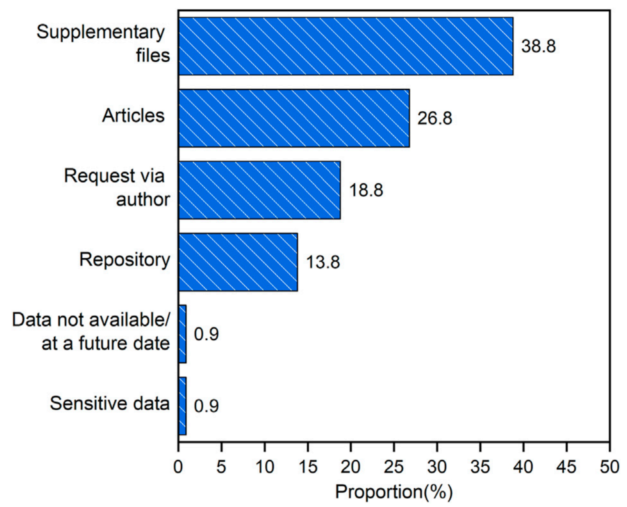

FIGURE 2 | Data dissemination methods indicated in the data availability statements of the articles analysed in this study.

request the data ($n = 42$; 18.8%). Data were explicitly shared via a data repository in only 31 (13.8%) articles.

Microplastics Data Sharing Trends in Data Repositories

In our sample set derived from Web of Science, the earliest microplastic article was published in 2006 while our data repository search yielded the first dataset in 2013 hence in this section our analysis is focused from 2013 to 2021 (**Figure 3**). Searches in the google dataset, dataONE Data One (2015) and OpenAIRE discovery portals returned 72 datasets on microplastics. Further site-specific searches of 10 data repositories increased the number of datasets to 279 (**Table 2**; **Figure 4**). Of these 279 datasets, 222 (79.6%) had data files that were directly accessible, while for 57 datasets (20.4%) the files were not accessible (*i.e.*, only metadata were provided) or further approval was required to access and download the data files. Search results were not limited by year, but the first datasets in the sample were published in the year 2013, with evidence of data sharing in data repositories trending up thereafter (**Figure 3A**).

In addition to the datasets in repositories, 6,363 microplastics articles were published between the years 2013 and 2021 (**Figure 3B**). During this time, the number of articles increased exponentially from 50 to 2,295, while the number of datasets from those studies provided within the repositories also increased rapidly from 1 to 112 (**Figure 4**). Of the 10 repositories queried, the CoreTrustSeal certified repositories Pangaea and Mendeley published the majority of microplastics datasets, with approximately 32 and 27%, respectively.

Geographic Distribution of Data Site Locations

The metadata collected highlights the unequal geographic distribution of the provenance of the microplastics samples (**Figure 5**). Of the 279 repository datasets, the majority of data were sampled in Europe (13.4%), North America (12.8%), and in

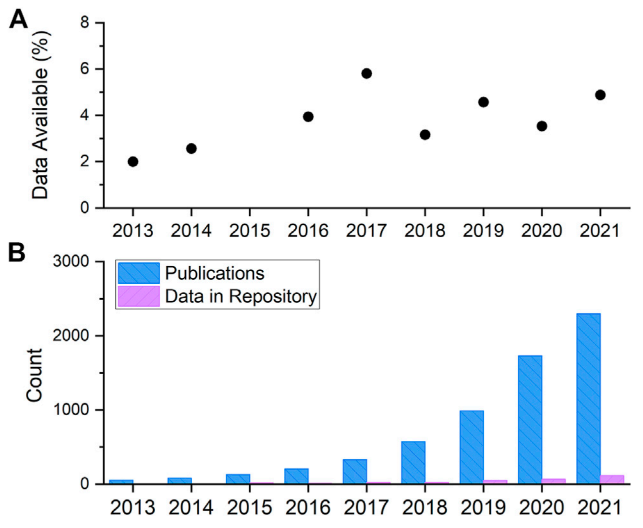

FIGURE 3 | (A) Proportion of microplastic publications found in Web of Science between 2013 and 2021 that have a dataset in an open access repository; **(B)** Number of data sets in open access repositories during the same period along with the annual number of peer-reviewed microplastics articles published.

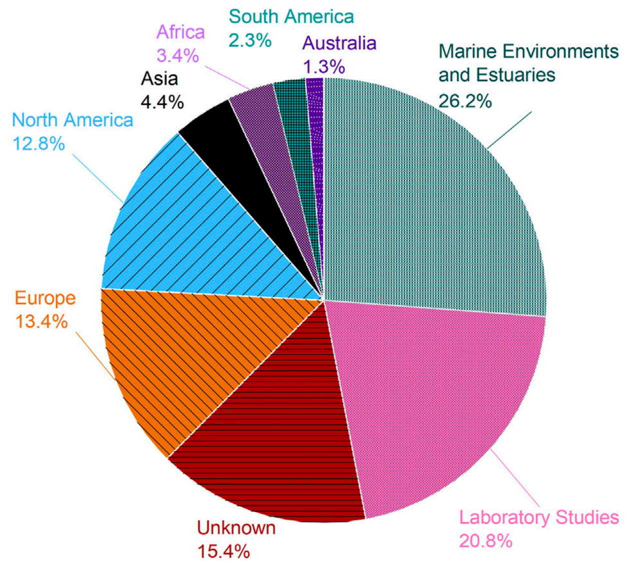

FIGURE 5 | Geographical distribution of microplastics datasets according to sample provenance. Note that the marine environments and estuaries section includes beaches, coastal areas, seas, and trenches. Unknown means that there were not sufficient metadata in the dataset to determine the location.

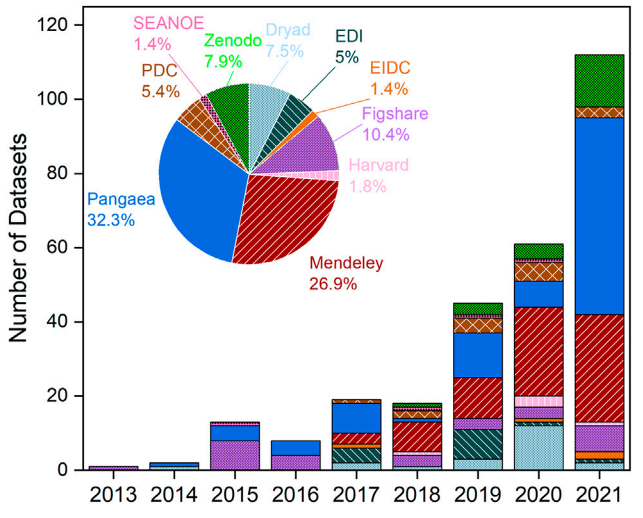

FIGURE 4 | Yearly (bar graph) and cumulative (pie chart) distributions of microplastics datasets uploaded to open access repositories between 2013 and 2021.

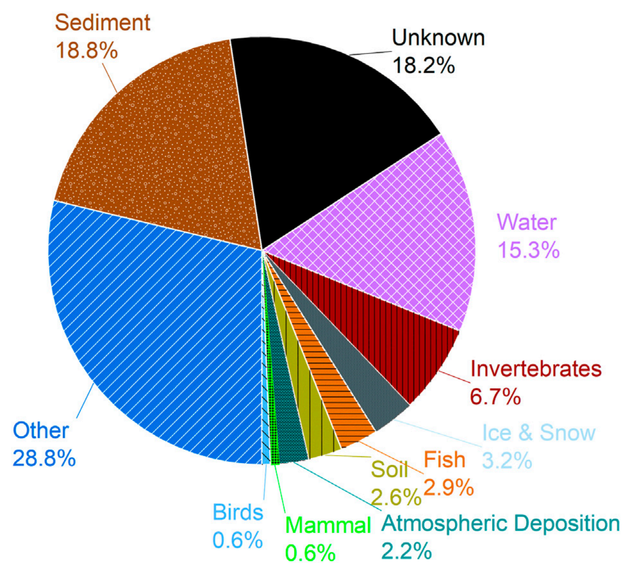

FIGURE 6 | Distribution of open datasets by media type. The "other" section includes effect and fate studies, as well as methods papers and modelling experiments. Unknown means that there were not sufficient metadata in the dataset to determine the media type.

marine environments and estuaries (26.2%). Approximately 11% of the data originated from Australia, South America, Asia, plus Africa. About a fifth of the datasets were generated in controlled laboratory studies. The latter include, for example, studies looking at uptake of microplastic particles by biological organisms or microplastic particle transport in porous media (20.8%), while 46 studies (15.4%) did not report any location information as part of the repository dataset.

Environmental Media Type Reported With Dataset

The largest fraction of datasets (28.8%) contained data from studies that focused on purchased plastics, uptake of microplastics in organisms, and modelling experiments. The most reported media types included sediments (18.8%), water (15.3%), and invertebrates (6.7%), with other types of studies accounting for the remaining 12.1%. The least studied media were atmospheric deposition, mammals, the cryosphere (ice and snow), fish, and birds. An unexpected 18.2% of datasets did not provide sufficient information to identify the media type (**Figure 6**).

DISCUSSION

Although the number of new datasets made available publicly are increasing annually, the numbers continue to be very low relative to the rapid growth in articles about environmental microplastics (**Figures 1**, **3**, **4**). The increase in data sharing, especially since 2016, are likely because publishers have implemented data policies which state that authors are expected to include an explicit data sharing statement within their manuscript at the time of submission (e.g., Piwowar and Vision, 2013; Science, 2019; Colavizza et al., 2020; AGU, 2021; Elsevier, 2022; Springer, 2022). Less than 30% of microplastic research articles assessed in this study included any form of data sharing statement, with only 13.8% of articles explicitly sharing their data in a repository. Pangaea and Mendeley were among the most commonly used repositories, perhaps because they are free and easy to use (**Figure 4**). It is not entirely known why these were most used, but the earth science Pangaea community has a long history of depositing and archiving data. In addition, numerous microplastic researchers are based at one of the host institutes of PANGAEA (AWI), which encouraged sustainable data archiving early on and often curated data produced in large European Union projects where it was a partner. Mendeley Data was purchased in 2013 by Elsevier and researchers publishing in Elsevier journals are encouraged to deposit their data in Mendeley Data (Dumon, 2013). It is possible that many of these researchers are unaware of other repository options and, hence, may gravitate towards using the publisher's controlled repositories.

The challenge of finding and accessing research data is not unique to the microplastics research community. Similar patterns in data sharing practices are common in well-established disciplines such as social sciences, water resources, low-temperature geochemistry, ecology, and health sciences (Stagge et al., 2019; Brantley et al., 2021; Tedersoo et al., 2021; Roche et al., 2022b). For example, Stagge et al. (2019) assessed data availability and research reproducibility in hydrology and water resources across several journals and found that, while approximately 70% of the sampled articles stated some materials were available, only around 48% of the materials could be accessed online. In experimental biology, only one in five papers (21.5%) included a data sharing statement or associated open data (Roche et al., 2022b). An overview of published research funded by the Canadian Institutes of Health Research (CIHR) (Government of Canada, 2021), a federal funding agency which has an explicit data sharing expectation, showed that for a subset of CIHR funded projects, only 45.2% of studies had readily accessible data (Read et al., 2021). The challenges outlined by Brantley et al. (2021) for the field of Earth surface geochemistry similarly resonated with our assessment of the emerging field of environmental microplastics. Perhaps the biggest challenge faced in both cases is the diverse nature of the data due to the environmental media involved, QA/QC issues, data structure, diversity in analytical techniques used, and multiple other factors, which makes it challenging to develop standardised reporting structures (Cowger et al., 2020; Provencher et al., 2020; Brandes et al., 2021; Brantley et al., 2021). A promising trend in all the disciplines mentioned, microplastics included, is that more and more researchers are making their data available. In addition to publishers' data policies, this increase may be attributed to the generational shift with the research community as younger researchers are getting more access to technology and databases, and early exposure to the data management concepts and practices; they are integrating all of these as part of their daily research workflow.

Microplastics are a relatively young field, and thus it can be expected that it will lag behind more established disciplines with respect to data sharing, especially given its multi-disciplinary nature. In this regard, researchers are likely to be influenced by their home discipline which may slow consensus on the discipline-specific metadata and data sharing standards, guidance, and education. However, the microplastics data sharing practices observed in this study showed that the microplastics field is on par with well-established disciplines such as water resources and ecology (Stagge et al., 2019; Roche et al., 2022a). Efforts to develop microplastic metadata sharing practices, which will increase the findability and interoperability of microplastics data, are currently underway (Cowger et al., 2020; Cowger et al., 2020; AMAP, 2021; Jenkins et al., 2021). Such methods or other regulatory measures and incentives are urgently needed because the progress in data sharing over the past 3 years appears to have stabilised instead of continuing on an upward trajectory. Given the early stages of this area of research, these valuable data are not easily discoverable via peer reviewed literature and data repositories. However, they often constitute vitally important baselines needed for future monitoring purposes. As the data collection efforts expand to include indigenous lands in North America, data management should additionally adhere to guiding principles for data collected on indigenous lands such as the Collective Benefit, Authority to Control, and Ethics (CARE; Carroll et al., 2020; Carroll et al., 2021) and the Ownership, Control, Access and Possession (OCAP®) (FNIGC, 2020).

The microplastic research community can learn from, and lean on work in other disciplines to promote good practices for data sharing. There are a growing number of research data management best practice guidance papers available (Michener, 2015; Wilkinson et al., 2016; Briney et al., 2020; Persaud et al., 2021, Contaxis et al., 2022, among others). As emphasised by Brantley et al. (2021), targeted education and awareness are still needed across scientific disciplines in order to implement and sustain best data management practices. Given the rapid growth of microplastic research papers, the microplastic research community, the target audience of this paper, is unlikely to have the time to thoroughly review existing papers that have been published about research data management (RDM) standards and best practices for other fields that are transferable to the field of environmental microplastics. Five simple strategies for advancing good data management and data sharing practices in microplastics research are therefore provided in the next section. We hope that these will help to maximise the positive impacts of microplastic research and improve the FAIRness of microplastic research data.

Strategies for Advancing Good Research Data Management Practices in Microplastics Research

1) Use Available Standards/Practices to Describe Data

A major challenge in translating the rapidly increasing body of new scientific knowledge and data into actionable policy is the lack of standardised procedures for microplastics RDM practices. There are currently no international or national data governance standards for environmental microplastics, including metadata standards, database structures, and RDM best practices, which limits the effective sharing and comparison of data on the abundance, size distribution, shape, surface roughness and chemical (polymer) composition of microplastics. This, in turn, hampers efforts to harmonise, and eventually standardise, the evaluation and validation of sampling and analytical methodologies and protocols that are needed across the research community. This study acknowledges there are many challenges that still need to be addressed to standardise data reporting for microplastics research, however, resources such as the AMAP report (2021), GESAMP (2019), Cowger et al. (2020), Michida et al. (2020), Jenkins et al. (2021) and Miller et al. (2021) provide guidelines that will help ensure data collection and reporting are robust. Existing metadata standards, such as the United States EPA Water Quality Exchange, the Dublin Core™ Metadata Element Set, and European monitoring under the Marine Strategy Framework Directive (EMODnet) should be integrated to ensure data are described consistently across the microplastics community.

2) Share Raw Data - Or as Close to Raw as Possible

It may be necessary to perform QA/QC, or to transform data from a format that is ideal for analysis into a format that is ideal for accessibility (*e.g.*, CSV, mzML, JCAMP-DX, JSON, cif, TIF), but the goal should always be to share data that are as close to raw as possible. The dataset should include a README with information on how, when, and where data were collected and any pre- and post-processing steps, which travel alongside the data and provide necessary context and contact information. If any data were provided by a third-party source, or derived from data provided by a third-party, that information should be documented with the dataset.

3) Use a Trusted Digital Repository

Whenever possible, data should be shared in a trusted digital repository that will steward data in the long term. Ideally, the repository will provide DOIs or another unique and persistent identifier that can be used to reference individual datasets. Embargoes may be used to temporarily protect data from downloads, especially if analyses are still ongoing. Some repositories can also restrict access to data in the longer term if they are sensitive. Disciplinary repositories, such as NOAA's NCEI Microplastics database, accept marine microplastics data from all researchers across the globe (NCEI Microplastics nd), while specifically within the European Union member states as a requirement for EMODNet. Otherwise, researchers are encouraged to use general-purpose or institutional repositories, such as Pangaea and the Federated Research Data Repository (FRDR), which offer curation services to deposited data.

4) Link Dataset to Publications

Many journals provide either supplemental information or data availability/open access statements, where the repository name and the dataset DOI (or other identifiers) should be included so readers can find the supporting data. Data obtained from a third-party source should also be included and cited in the references section. If the data cannot be shared, or if restricted access is required, this should be explicitly stated and the steps required to obtain access outlined. Likewise, the DOI of any publications that is associated with the dataset can be added to the data repository metadata record which provides context for the data and positions it as an important part of the scholarly record.

5) Plan to Share Data From the Onset of a Study

Data management and data sharing should be considered as early as possible. A data management plan can be completed at any point in the research process to document what types of data are generated, their format, and the metadata standards that are used to describe them, as well as short and long-term storage requirements, and the costs associated with data collection and data management. Planning early on helps ensure no data are lost and that the resulting dataset relies on existing practices to ensure a measure of consistency and interoperability, and, when necessary, that permission to share data has been sought and provided (*e.g.*, for data that were provided by a third-party, or data that were collected on Indigenous lands or with Indigenous partners).

CONCLUSION

In this bibliometric study, the extent to which environmental microplastic research data are openly shared were assessed. This work showed that between 2013 and 2021, microplastics dataset sharing has increased, but much more slowly than the number of peer-reviewed publications. The large amounts of data being produced, often supported by public funding, are simply not accessible or have insufficient metadata for others to do quality assurance, to assess the quality of the data and to ultimately reuse the data. Data sharing has stabilised in recent years which suggests that there are obstacles to data sharing that will need to be addressed to ensure the long-term availability and accessibility of data which can serve as vital baseline data for future monitoring. For example, many institutions need to access microplastics data to help guide regulatory frameworks such as safe drinking water levels (California Senate Bill 1422, 2018), and the European Marine Strategy Framework Directive (European European Union, 2008). These findings highlight the need for the environmental microplastics community to focus on not only advancing the science of environmental microplastics research but also on simultaneously embedding data management into their daily research workflow through education and best practices. Efforts should be made by researchers to also make use of data management resources including sharing data on discipline-specific repositories that are available to ingest microplastics data. However, standardised (meta)data reporting templates that implement established microplastic (meta)data reporting standards in a reproducible and usable way for microplastics data are still needed.

The increasing trend of open microplastic data shared in repositories and linked to peer-reviewed publications is promising. Data sharing practices will help increase the reproducibility and comparability of data. The more comprehensive our collective data sharing practices are, the better the microplastics decisions and policies that affect society as a whole. Moreover, it is incumbent upon the microplastics research community to ensure that FAIR data are consistently made available. These types of activities will not only strengthen the data sharing practices in this field but will also support continued advances in understanding the occurrence of microplastics in the environment, which is highly important in terms of pollution monitoring efforts.

DATA AVAILABILITY STATEMENT

The datasets generated and analysed in this study can be found in the Federated Research Data Repository at https://doi.org/10.20383/102.0476.

AUTHOR CONTRIBUTIONS

All co authors were involved in conceptualization of study, data collection, reviewed and edited the manuscript. All authors read and approved the final manuscript.

FUNDING

This work was primarily funded by the NSERC/ECCC Alliance Grants - Plastics science for a cleaner future program (Grant Agreement #: ALLRP 558435-20). Additional funding support is also acknowledged from the following: the Canada First Research Excellence Fund Global Water Futures Programme, the McPike Zima Charitable, the PoF IV program "Changing Earth - Sustaining our Future" Topic 6.4 of the German Helmholtz Association, the Research Council of Norway projects REVEAL (Grant agreement #. 301157), ANDROMEDA (Grant agreement # 312262), MicroLEACH (Grant agreement #. 295174), the Early-Career Research Fellowship from the Gulf Research Program of the US National Academies of Sciences, Engineering, and Medicine (Grant agreement # 2000012639), European Union's Horizon 2020 Coordination and Support Action programme under Grant agreement 101003805 (EUROqCHARM), and the Union Horizon 2020 research and innovation programme under Marie Skłodowska-Curie grant agreement no. 838237-OPTIMISE.

ACKNOWLEDGMENTS

This publication is Eprint ID 55791 of the Alfred-Wegener-Institut Helmholtz-Zentrum für Polar-und Meeresforschung. We acknowledge Dr. Martin Wagner's inputs during project conceptualization and on an earlier draft of manuscript from Norwegian University of Science and Technology. We also thank Mickey Nielsen at the University of Waterloo for her project coordination support.

REFERENCES

Abeynayaka, A., and Norihiro, I. (2019). "A Framework to Incorporate Aquatic Plastic into Life Cycle Assessment of Plastic Products," in *EcoDesing 2019 International Symposium* Yokohama, 261–265.

AGU (2021). Data and Software for Authors. Available at: https://www.agu.org/Publish-with-AGU/Publish/Author-Resources/Data-and-Software-for-Authors (Accessed, 2022).

AMAP (2021). *Overview of AMAP Initiatives of Monitoring and Assessment of Plastic Pollution in the Arctic*. AMAP: AMAP. Available at: https://www.amap.no/documents/download/6714/inline (Accessed, 2022).

Barboza, L. G. A., Dick Vethaak, A., Lavorante, B. R. B. O., Lundebye, A.-K., and Guilhermino, L. (2018). Marine Microplastic Debris: An Emerging Issue for Food Security, Food Safety and Human Health. *Mar. Pollut. Bull.* 133, 336–348. doi:10.1016/j.marpolbul.2018.05.047

Bergmann, M., Gutow, L., and Klages, M. (2015). *Marine Anthropogenic Litter*. Laporte: Springer Cham. doi:10.1007/978-3-319-16510-3

Brandes, E., Henseler, M., and Kreins, P. (2021). Identifying Hot-Spots for Microplastic Contamination in Agricultural Soils-A Spatial Modelling Approach for Germany. *Environ. Res. Lett.* 16, 104041. doi:10.1088/1748-9326/ac21e6

Brantley, S. L., Wen, T., Agarwal, D. A., Catalano, J. G., Schroeder, P. A., Lehnert, K., et al. (2021). The Future Low-Temperature Geochemical Data-Scape as Envisioned by the U.S. Geochemical Community. *Comput. Geosciences* 157, 104933. doi:10.1016/j.cageo.2021.104933

Briney, K. A., Coates, H., and Goben, A. (2020). Foundational Practices of Research Data Management. *Res. Ideas Outcomes* 6, e56508. doi:10.3897/rio.6.e56508

California Senate Bill (Sb) 1422 (2018). California Safe Drinking Water Act: Microplastics. Available at: https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201720180SB1422.

Can-Güven, E. (2021). Microplastics as Emerging Atmospheric Pollutants: A Review and Bibliometric Analysis. *Air Qual. Atmos. Health* 14, 203–215. doi:10.1007/s11869-020-00926-3

Carroll, S. R., Garba, I., Figueroa-Rodríguez, O. L., Holbrook, J., Lovett, R., Materechera, S., et al. (2020). The CARE Principles for Indigenous Data Governance. *Misc. Community Pract.* 19, 1. doi:10.5334/dsj-2020-043

Carroll, S. R., Herczog, E., Hudson, M., Russell, K., and Stall, S. (2021). Operationalizing the CARE and FAIR Principles for Indigenous Data Futures. *Sci. Data* 8, 108. doi:10.1038/s41597-021-00892-0

Colavizza, G., Hrynaszkiewicz, I., Staden, I., Whitaker, K., and McGillivray, B. (2020). The Citation Advantage of Linking Publications to Research Data. *PLoS ONE* 15, e0230416. doi:10.1371/journal.pone.0230416

Contaxis, N., Clark, J., Dellureficio, A., Gonzales, S., Mannheimer, S., Oxley, P. R., et al. (2022). Ten Simple Rules for Improving Research Data Discovery. *PLoS Comput. Biol.* 18, e1009768. doi:10.1371/journal.pcbi.1009768

CoreTrustSeal (2022). Core Trust Seal. Available at: https://www.coretrustseal.org (Accessed, 2022).

Covernton, G. A., Pearce, C. M., Gurney-Smith, H. J., Chastain, S. G., Ross, P. S., Dower, J. F., et al. (2019). Size and Shape Matter: A Preliminary Analysis of Microplastic Sampling Technique in Seawater Studies with Implications for Ecological Risk Assessment. *Sci. Total Environ.* 667, 124–132. doi:10.1016/j.scitotenv.2019.02.346

Cowger, W., Booth, A. M., Hamilton, B. M., Thaysen, C., Primpke, S., Munno, K., et al. (2020). Reporting Guidelines to Increase the Reproducibility and Comparability of Research on Microplastics. *Appl. Spectrosc.* 74, 1066–1077. doi:10.1177/0003702820930292

Data One (2015). Data One. Available at: https://www.dataone.org (Accessed, 2021).

Domercq, P., Praetorius, A., and MacLeod, M. (2022). The Full Multi: An Open-Source Framework for Modelling the Transport and Fate of Nano- and Microplastics in Aquatic Systems. *Environ. Model. Softw.* 148, 105291. doi:10.1016/j.envsoft.2021.105291

Dumon, O. (2013). Elsevier Welcomes Mendeley. Available at: https://www.elsevier.com/connect/elsevier-welcomes-mendeley (Accessed, 2022).

Elsevier (2022). Research Data. Available at: https://www.elsevier.com/about/policies/research-data (Accessed, 2022).

Environment CanadaClimate Change Canada (2019). Canada's Plastics Science Agenda. Government of Canada. Available at: https://www.canada.ca/en/environment-climate-change/services/science-technology/canada-science-plastic-agenda.html (Accessed, 2022).

FNIGC (2020). The First Nations Principles of OCAP. Available at: https://fnigc.ca/ocap-training/ (Accessed, 2022).

GESAMP (2019). Reports and Studies Guidelines for the Monitoring and Assessment of Plastic Litter in the Ocean. Available at: http://www.gesamp.org/publications/guidelines-for-the-monitoring-and-assessment-of-plastic-litter-in-the-ocean (Accessed, 2022).

Government of Canada (2021). Tri-Agency Research Data Management Policy. Ottawa, ON: Innovation, Science and Economic Development Canada. Available at: https://www.science.gc.ca/eic/site/063.nsf/eng/h_97610.html (Accessed, 2022).

Igalavithana, A. D., Mahagamage, M. G. Y. L., Gajanayake, P., Abeynayaka, A., Gamaralalage, P. J. D., Ohgaki, M., et al. (2022). Microplastics and Potentially Toxic Elements: Potential Human Exposure Pathways Through Agricultural Lands and Policy Based Countermeasures. *Microplastics* 1 (1), 102–120. doi:10.3390/microplastics1010007

Jacques, O., and Prosser, R. S. (2021). A Probabilistic Risk Assessment of Microplastics in Soil Ecosystems. *Sci. Total Environ.* 757, 143987. doi:10.1016/j.scitotenv.2020.143987

Jenkins, T., Smith, R., Goucher, N., Persaud, B., Slowinski, S., Szigeti, K., et al. (2021). Workshop Report: Maximizing the Value of Environmental Microplastics Data. *Zenodo*, 4–13. doi:10.5281/zenodo.5710744

Koelmans, A. A., Mohamed Nor, N. H., Hermsen, E., Kooi, M., Mintenig, S. M., and De France, J. (2019). Microplastics in Freshwaters and Drinking Water: Critical Review and Assessment of Data Quality. *Water Res.* 155, 410–422. doi:10.1016/j.watres.2019.02.054

Li, J., Liu, H., and Paul Chen, J. (2018). Microplastics in Freshwater Systems: A Review on Occurrence, Environmental Effects, and Methods for Microplastics Detection. *Water Res.* 137, 362–374. doi:10.1016/j.watres.2017.12.056

Maes, T., Perry, J., Alliji, K., Clarke, C., and Birchenough, S. N. R. (2019). Shades of Grey: Marine Litter Research Developments in Europe. *Mar. Pollut. Bull.* 146, 274–281. doi:10.1016/j.marpolbul.2019.06.019

Michener, W. K. (2015). Ten Simple Rules for Creating a Good Data Management Plan. *PLoS Comput. Biol.* 11, e1004525. doi:10.1371/journal.pcbi.1004525

Michida, Y. (2020). *Guidelines for Harmonizing Ocean Surface Microplastic Monitoring Methods*. Tokyo, Japan: Chiyoda-ku Ministry of the Environment Japan, 71. Version 1.1. doi:10.25607/OBP-867

Miller, E., Sedlak, M., Lin, D., Box, C., Holleman, C., Rochman, C. M., et al. (2021). Recommended Best Practices for Collecting, Analyzing, and Reporting Microplastics in Environmental Media: Lessons Learned from Comprehensive Monitoring of San Francisco Bay. *J. Hazard. Mater.* 409, 124770. doi:10.1016/j.jhazmat.2020.124770

NOAA and NECEI (2022). NCEI Marine Microplastic. Available at: https://www.ncei.noaa.gov/maps/microplastics/ (Accessed, 2022).

NSERC (2020). Plastics Science for a Cleaner Future. Available at: https://www.nserc-crsng.gc.ca/professors-professeurs/rpp-pp/plastics-plastiques_eng.asp (Accessed, 2022).

OpenAIRE (2013).. Available at: https://www.openaire.eu (Accessed, 2022).

Persaud, B. D., Dukacz, K. A., Saha, G. C., Peterson, A., Moradi, L., O'Hearn, S., et al. (2021). Ten Best Practices to Strengthen Stewardship and Sharing of Water Science Data in Canada. *Hydrol. Process.* 35, e14385. doi:10.1002/hyp.14385

Persson, L., Carney Almroth, B. M., Collins, C. D., Cornell, S., de Wit, C. A., Diamond, M. L., et al. (2022). Outside the Safe Operating Space of the Planetary Boundary for Novel Entities. *Environ. Sci. Technol.* 56, 1510–1521. doi:10.1021/acs.est.1c04158

Piwowar, H. A., and Vision, T. J. (2013). Data Reuse and the Open Data Citation Advantage. PeerJ 1, e175. doi:10.7717/peerj.175

Provencher, J. F., Covernton, G. A., Moore, R. C., Horn, D. A., Conkle, J. L., and Lusher, A. L. (2020). Proceed with Caution: The Need to Raise the Publication Bar for Microplastics Research. *Sci. Total Environ.* 748, 141426. doi:10.1016/j.scitotenv.2020.141426

Read, K. B., Ganshorn, H., Rutley, S., and Scott, D. R. (2021). Data-Sharing Practices in Publications Funded by the Canadian Institutes of Health Research: A Descriptive Analysis. *cmajo* 9, E980–E987. doi:10.9778/cmajo.20200303

Roche, D. G., Berberi, I., Dhane, F., Lauzon, F., Soeharjono, S., Dakin, R., et al. (2022a). Slow Improvement to the Archiving Quality of Open Datasets Shared by Researchers in Ecology and Evolution. *Proc. R. Soc. B.* 289, 1975. doi:10.1098/rspb.2021.2780

Roche, D. G., Raby, G. D., Norin, T., Ern, R., Scheuffele, H., Skeeles, M., et al. (2022b). Paths Towards Greater Consensus Building in Experimental Biology. *Exp. Biol.* 225, jeb243559. doi:10.1242/jeb.243559

Ryan, P. G. (2015). "A Brief History of Marine Litter Research,". Editors M. Bergmann, L. Gutow, and M. Klages (Cham: Springer International Publishing), 1–25. Mar. Anthropog. Litter. doi:10.1007/978-3-319-16510-3_1

Science (2019). Science Journals: Editorial Process. Available at: https://www.science.org/content/page/science-journals-editorial-policies#top (Accessed, 2022).

Springer (2022). Data Availability Statement. Available at: https://www.springer.com/gp/editorial-policies/data-availability-statement (Accessed, 2022).

Stagge, J. H., Rosenberg, D. E., Abdallah, A. M., Akbar, H., Attallah, N. A., and James, R. (2019). Assessing Data Availability and Research Reproducibility in Hydrology and Water Resources. *Sci. Data* 6, 190030. doi:10.1038/sdata.2019.30

Morgan Stanley, National Geographic, University of Georgia, and National Oceanic Atmospheric Administration (2010). Marine Debris Tracker. Available at: https://debristracker.org (Accessed, 2022).

State of California (2022). California Takes Decisive Action to Reduce Microplastics Pollution: State Adopts a First-In-Nation Approach to Protecting Ocean and Human Health. Available at: https://www.opc.ca.gov/2022/02/california-takes-decisive-action-to-reduce-microplastics-pollution-state-adopts-a-first-in-nation-approach-to-protecting-ocean-and-human-health/(Accessed, 2022).

Stokstad, E. (2022). World's Nations Start to Hammer Out First Global Treaty on Plastic Pollution. *News From Science*. [Online]. (Last Updated 12:55 PM on 23 February 2022) (Accessed March 9, 2022).

Tedersoo, L., Küngas, R., Oras, E., Köster, K., Eenmaa, H., Leijen, Ä., et al. (2021). Data Sharing Practices and Data Availability Upon Request Differ Across Scientific Disciplines. *Sci. Data* 8, 192. doi:10.1038/s41597-021-00981-0

Tekman, M. B., Gutow, L., Bergmann, M., and Peter, C. (2020). Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research and Meeresforschung. LITTERBASE. Online Portal for Marine Litter. Available at: https://litterbase.awi.de (Accessed, 2022).

M. B. Tekman, B. A. Walther, C. Peter, L. Gutow, and M. Bergmann, 2022. *Impacts of Plastic Pollution in the Oceans on Marine Species, Biodiversity and Ecosystems*. Berlin: WWF Germany. doi:10.5281/zenodo.5898684

European Union (2008). The Marine Strategy Framework Directive. Available at: https://ec.europa.eu/environment/marine/eu-coast-and-marine-policy/marine-strategy-framework-directive/index_en.htm.

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., et al. (2016). The FAIR Guiding Principles for Scientific Data Management and Stewardship. *Sci. Data* 3, 160018. doi:10.1038/sdata.2016.18

Woods, J. S., Verones, F., Jolliet, O., Vázquez-Rowe, I., and Boulay, A.-M. (2021). A Framework for the Assessment of Marine Litter Impacts in Life Cycle Impact Assessment. *Ecol. Indic.* 129, 107918. doi:10.1016/j.ecolind.2021.107918

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.