

RESEARCH ARTICLE

Open Access



Trash Taxonomy Tool: harmonizing classification systems used to describe trash in environments

Hannah Hapich¹ , Win Cowger^{1,2*} , Andrew Gray¹ , Neil Tangri³ , Tony Hale⁴ , Amr Magdy⁵ , Antoinette Vermilye⁶, Walter Yu⁷, Dick Ayres⁸, Charles Moore^{2,9}, John Vermilye⁶, Samiksha Singh¹, Aaron N. K. Haiman¹⁰, Kathryn Youngblood¹¹, Yunfan Kang⁵, Margaret McCauley¹², Trevor Lok¹³, Shelly Moore^{2,3} , Eric Baggs¹⁴, Sherry Lippiatt¹⁵ , Peter Kohler¹⁶ , Gary Conley¹⁷, Janna Taing¹⁸ and Jeremiah Mock¹⁹ 

Abstract

Despite global efforts to monitor, mitigate against, and prevent trash (mismanaged solid waste) pollution, no harmonized trash typology system has been widely adopted worldwide. This impedes the merging of datasets and comparative analyses. We addressed this problem by 1) assessing the state of trash typology and comparability, 2) developing a standardized and harmonized framework of relational tables and tools, and 3) informing practitioners about challenges and potential solutions. We analyzed 68 trash survey lists to assess similarities and differences in classification. We created comprehensive harmonized hierarchical tables and alias tables for item and material classes. On average, the 68 survey lists had 20.8% of item classes in common and 29.9% of material classes in common. Multiple correspondence analysis showed that the 68 surveys were not significantly different regarding organization type, ecosystem focus, or substrate focus. We built the Trash Taxonomy Tool (TTT) web-based application with query features and open access at openanalysis.org/trashtaxonomy. The TTT can be applied to improve, create, and compare trash surveys, and provides practitioners with tools to integrate datasets and maximize comparability. The use of TTT will ultimately facilitate improvements in assessing trends across space and time, identifying targets for mitigation, evaluating the effectiveness of prevention measures, informing policymaking, and holding producers responsible.

Keywords: Plastic, Pollution, Waste, Anthropogenic, Surveys, Harmonization, Characterization, Trash, Litter, Typology

Introduction

It is widely recognized that the impacts of mismanaged trash on ecosystems pose substantial risks to the environment. Mismanaged trash can kill wildlife via ingestion and entanglement [1] and transport invasive species [2, 3]. We use the term “trash” here and throughout this manuscript, meaning any anthropogenic object that has

escaped management and entered the environment. We recognize that in some localities, the term “litter,” “anthropogenic litter,” or “marine debris” may be more common synonyms, but we refrain from using it here, recognizing that not all trash comes from the act of littering, or is found in marine environments [4]. Trash on land has the potential to be transported through stormwater to other environments (e.g. rivers, the ocean) and contribute to persistent accumulation in the environment [3, 5]. In response to this growing threat, several communities have implemented water quality regulations [6, 7] to reduce trash delivery to aquatic water systems; a

*Correspondence: wincowger@gmail.com; wcowg001@ucr.edu

¹ Department of Environmental Science, University of California, 900 University Ave, Riverside, CA 92521, USA
Full list of author information is available at the end of the article

common way to comply with these regulations is through trash surveys. Surveys of trash in streams, beaches, oceans, and other environmental compartments are conducted to assess risks [8], plan mitigation [9], determine prevention priorities [10], and inform policymaking [11]. Trash surveys often include trash typology (e.g., bottle, plastic, cigarette), trash abundances, and site descriptions.

Trash typology systems are often developed with specific use cases or objectives in mind, which can hamper comparability. Trash survey developers try to minimize the number of classes they use to reduce complexity (and save time) while focusing on the classes most relevant to their management questions. For example, a trash survey (Survey 1) on a beach may explicitly list derelict *fishing gear* as a class. In contrast, another survey (Survey 2) may specify unique fishing items, such as *fishing poles and fishing wire*. In this example, *fishing gear* quantities from Survey 1 cannot be easily converted into *fishing poles* and *fishing wire* classes in Survey 2 without those being explicitly listed on the trash survey [12]. There has yet to be an in-depth analysis of the state of comparability between trash typologies.

Standardization and harmonization between existing trash surveys will allow trash monitoring data to be readily used together [13, 14]. Standardization involves prescribing one survey list or a set of survey lists for different use cases. The Joint Research Centre (JRC) and Oslo/Paris Convention (OSPAR) created standardized trash survey lists with regional and ecosystem foci which are in widespread use. Standardized trash typologies are also used to improve models, like image classification, by lessening image labeling labor and increasing class interpretability. Harmonization involves developing taxonomic frameworks that facilitate operations between established and new surveys. Hierarchical and alias frameworks (including relational tables) are the primary harmonization tools. Morales-Caselles et al. 2021 developed a spreadsheet for harmonizing multiple survey lists to a standardized format using aliases [15]. Vriend et al. [16] developed a framework for harmonizing river trash monitoring strategies, outlining six hierarchical levels of trash typology: 1) organic/inorganic, 2) material, 3) polymer classes, 4) polymer type, 5) item type, and 6) the raw sample. The JRC (Joint Research Centre) recently developed a hierarchical framework to allow heterogeneous surveys to be combined and analyzed with their standardized typology [17]. There has yet to be a framework proposed for complete harmonization and standardization of all trash survey typologies.

Our study aims to assess trash survey comparability and develop a framework to harmonize and standardize

trash surveys. To achieve these aims, we develop and describe the use of a universal trash typology, the Trash Taxonomy Tool (TTT). The TTT is a relational database (which can be queried by hierarchy and alias) and schema matching tool for harmonizing trash surveys. We use the TTT to assess the current state and future of trash typology, trash survey types, and survey comparability. We describe strategies to use the TTT to harmonize and standardize trash surveys. Recognizing that new trash objects are constantly being created, we discuss how to adapt and improve the TTT.

Methods

Developing relational tables

Approach and assumptions

We compiled 68 English-language survey lists from various countries and organizations, including government, research, nonprofit, and academic groups that describe trash survey types for freshwater, marine, and terrestrial ecosystems. Throughout this report, we italicize class names when referring to trash typology. Three groups of classes were found across most of the surveys, which describe trash in terms of materials (the resource used to make the item, e.g., *plastic* or *paper*), item (description of the form of the object, e.g., *bottle* or *fragment*), and brand (the logo or manufacturer's name identified on the item) (Fig. 1). We also recognized two relational systems within the data: alias (synonymous words, e.g., *cap* and *lid*) and hierarchy (words that are parents or nested as children, e.g., *spoon*, *fork*, and *knife* are nested under utensils). We developed relational tables for comparing words used within and between these structures that originated from the 68 selected survey sheets. To provide potential users with definitions we operationalized in this study, we present a glossary of terms used (Table S1).

The primary assumption within the TTT is that there are no differences in the definitions of a given class between surveys. An example of a violation of this assumption would be two surveys that define *fragment* based on size, but with different criteria: such as *fragment* = particles > 1 mm versus *fragment* = particles < 5 mm. These surveys would classify different sets of objects using the same word. There are other types of information held within the methodological distinctions in definitions that we did not investigate further (e.g., color, shape, size) unless the methodological limitation was encoded in the class name (e.g., rope diameter < 1 cm). This study compared the relationships between the words used to describe trash and how they relate to one another based on professional experience with trash nomenclature.

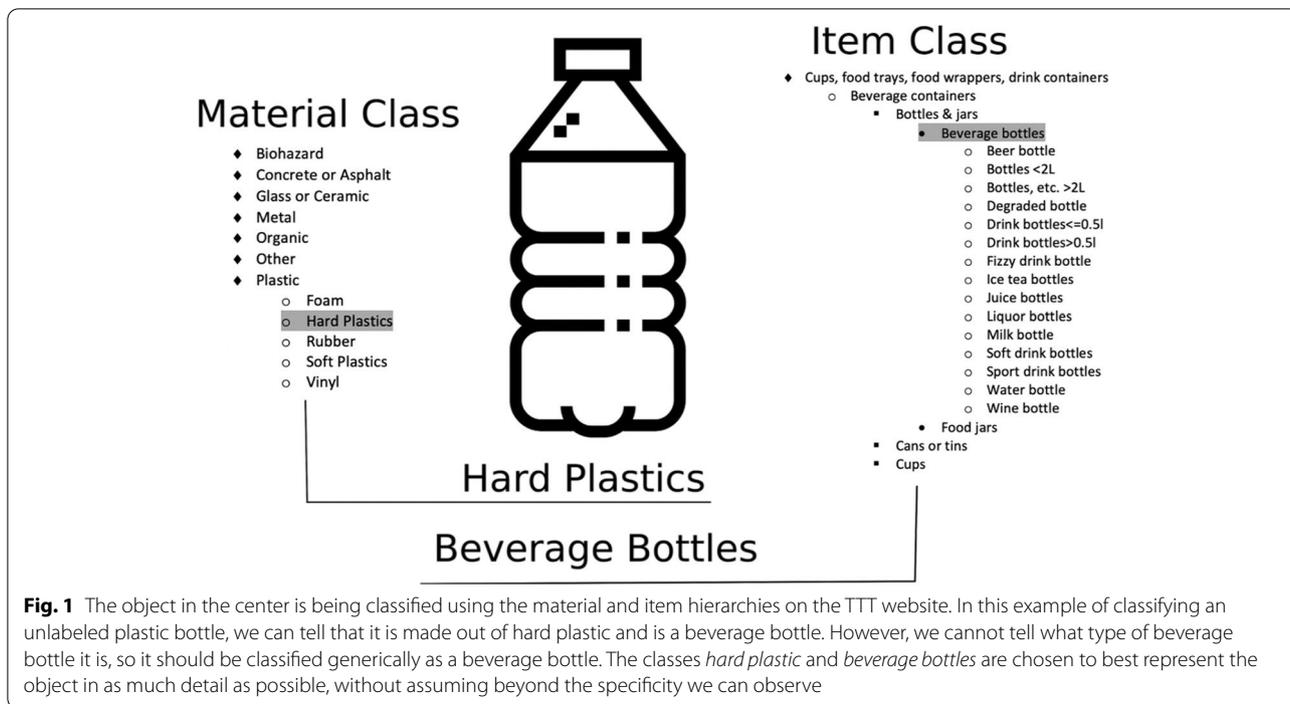


Fig. 1 The object in the center is being classified using the material and item hierarchies on the TTT website. In this example of classifying an unlabeled plastic bottle, we can tell that it is made out of hard plastic and is a beverage bottle. However, we cannot tell what type of beverage bottle it is, so it should be classified generically as a beverage bottle. The classes *hard plastic* and *beverage bottles* are chosen to best represent the object in as much detail as possible, without assuming beyond the specificity we can observe

Material-item relational table

We compiled a table listing the materials and items described by each organization’s classification system that we reviewed for our study (Fig. 2). Each row represents a unique material-item relationship (e.g., *plastic* and *straw* being listed in a row together). Sometimes it was unclear whether a class described a material class or an item class (e.g., *disposable fork*, typically made of plastic). To avoid introducing bias and adding words not used explicitly in the surveys, these classes were placed in the item class, and the material class was not inferred.

Misaligned category table

We defined misaligned classes as classes that did not fit within the material, item, or brand classes. If the class was too ambiguous, did not conform to the standard one-parent rule for hierarchical databases, or did not describe trash in environments, we added it to a separate document called the misaligned class table. Examples of misaligned classes include *construction materials*, *fishing gear*, and *tree*.

Alias and hierarchical tables

We developed alias tables for material, item, and brand classes independent of one another (Fig. 2). For the item and material alias tables, all words that were found to have the same meaning were linked using rows in a table where the first column defined the prime word, which was used as a key for joining to the hierarchy

(Fig. 2), while all other columns were defined as aliases (e.g., *fork* and *forks* will be under the same alias). Break Free From Plastic, a nonprofit organization promoting a global movement to create a future free from plastic pollution, developed the brand alias table by researching the manufacturers who own the brands found during their annual Brand Audit in 2018 and 2019 [12]. This table was formatted with recurring manufacturer classes in one column corresponding to each brand owned by that manufacturer. In the alias tables, prime words can be merged with the hierarchical tables and vice versa. We established a single alias rule for every word in the alias tables so that any word could only join to one prime word to simplify analysis procedures using the tables.

Additionally, we developed hierarchy tables for material classes (Fig. 3A) and item classes (Fig. 3B). These tables specify the hierarchical position of prime words through multi-level grouping (e.g., the *utensils* class encompasses *forks*, *knives*, *spoons*, and *straws*; *plastic* in materials includes *foam* and *soft plastic*). The hierarchy tables only describe the prime words from the alias tables since those words are equivalent to the other words used to describe trash. Hierarchical groups were sometimes obvious. For example, one survey we reviewed used the class *glass/ceramic* while another split the classes into *glass* and *ceramic*. In other cases, the relationships were more nuanced. For example, *organic* is a more general material description that includes materials like *wood* and *cloth*. We established a single parent rule for the

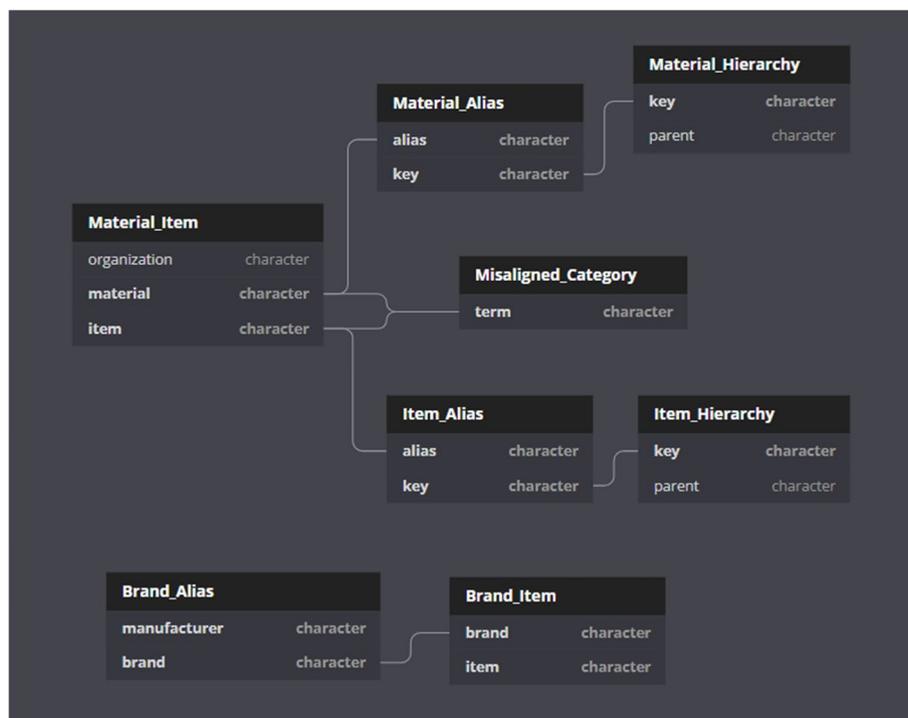


Fig. 2 Visual representation of relational tables and how they can be linked. Figure created with DB Diagram. This database can be accessed in an interactive version at <https://dbdiagram.io/d/5f3d9342cf48a141ff557dfe> where there is detailed information about each feature. Each box is a relational table. The rows are the column names in the table, followed by the information type (all are character type (i.e., text)). The lines show columns that are precisely equivalent in meaning and contents

hierarchical tables where every word could only have up to one parent to simplify analysis procedures using the tables.

Database query tool development

The Trash Taxonomy Tool (TTT) is a database with a set of query tools and all previously mentioned relational tables accessible via an online application (openanalysis.org/trashtaxonomy). The site was created using the shiny [18], dplyr [19], data.table [20], shinyjs [21], shinythemes [22], DT [23], shinyhelper [24], data.tree [25], and collapsibleTree [26], packages in R (4.0.5) and R Studio (1.4.1106). This site allows users to upload a comma-separated value (csv) file of their survey list to process using the alias and hierarchy framework; an example of the exact required formatting is provided in the supplemental information (Table S2). In technical terms, the TTT is a schema matching tool because it matches and maps schemas from trash surveys to a unified format [27]. The TTT first uses an alias lookup to match and map the user-provided survey classes to prime word keys for material and item words. It then locates the prime word in the hierarchy and allows users to display all recognized words that are more or less specific in their item

and material columns. It finds all parent words when the less specific function is called and all child words when the more specific function is called. If the user provides a word that is not in the relational tables, a notification will return for that particular word. More detailed documentation and a video tutorial (<https://youtu.be/sqeLaJKyol8>) can be found on the TTT website.

Relational table cleaning and validation

We cleaned the relational tables using several tests. We created basic queries to identify duplicated terms, remove them, and ensure that all relationship links between the tables (Fig. 2) were equivalent in both directions for the alias to hierarchy relationships. The material-item table keys are equivalent to the alias and misaligned prime keys combined. The key column in the 'material alias' table has the same terms as the key column in the 'material hierarchy' table. We created a visualization within our online tool to inspect all the relational tables for nuanced relationships like semantic relationships within and between the tables. We also uploaded the material-item relational table to the query tool, then returned the relational table's results and visually assessed the matches.

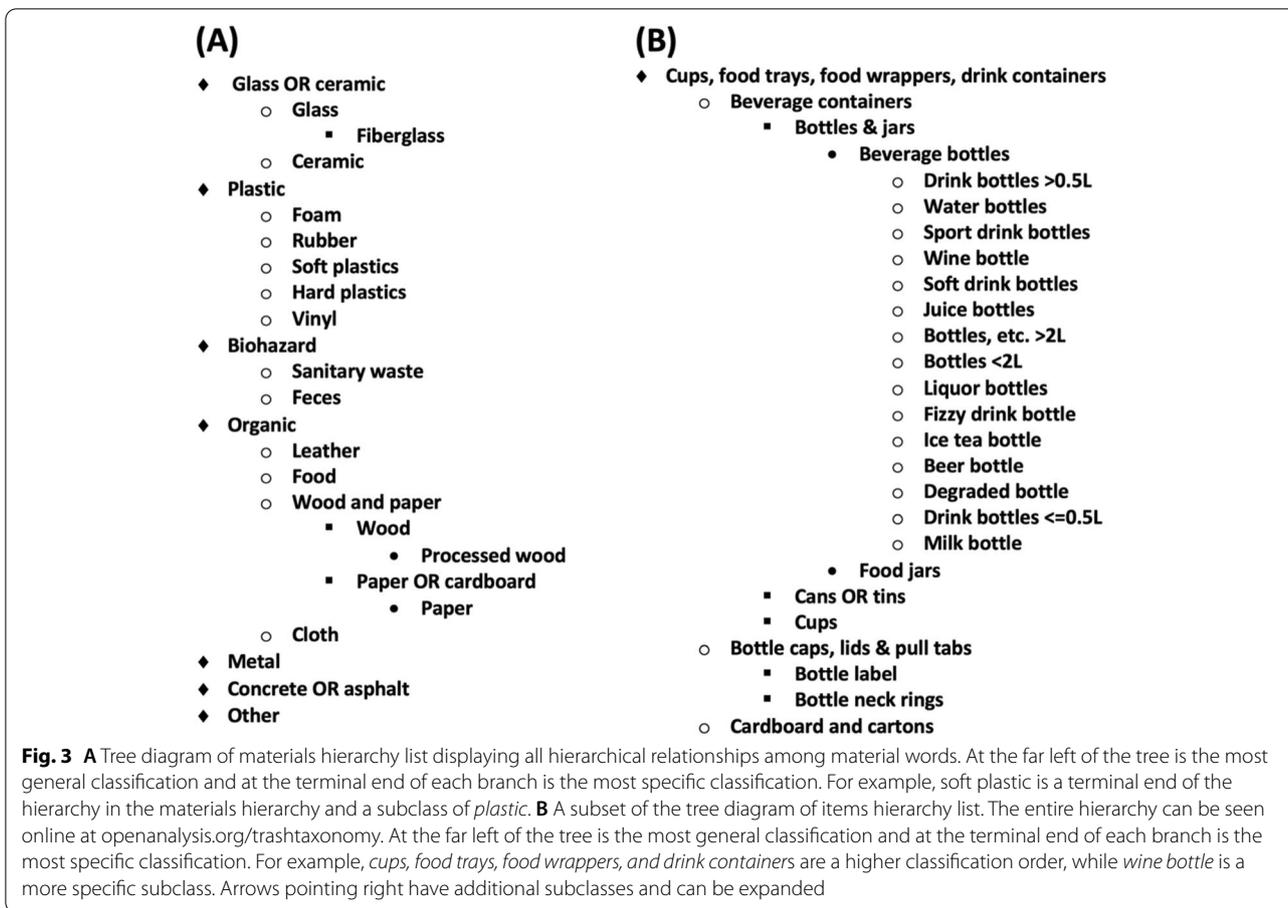


Fig. 3 **A** Tree diagram of materials hierarchy list displaying all hierarchical relationships among material words. At the far left of the tree is the most general classification and at the terminal end of each branch is the most specific classification. For example, soft plastic is a terminal end of the hierarchy in the materials hierarchy and a subclass of *plastic*. **B** A subset of the tree diagram of items hierarchy list. The entire hierarchy can be seen online at openanalysis.org/trashtaxonomy. At the far left of the tree is the most general classification and at the terminal end of each branch is the most specific classification. For example, *cups, food trays, food wrappers, and drink containers* are a higher classification order, while *wine bottle* is a more specific subclass. Arrows pointing right have additional subclasses and can be expanded

Assessment of the current state of trash typology

Summary statistics

We calculated summary statistics on each of the relational tables. The total number of classes was assessed by summing the number of unique words used within the survey lists (e.g., *fork or spoon* and *fork/spoon* are considered separate words). We assessed the number of unique classes by summing the unique prime aliases in the alias table (e.g., the two previously mentioned categories are joined to the same class). The number of levels of the hierarchies was assessed using the maximum number of levels of any given branch in the hierarchy tables. Diagrams were developed to demonstrate the depth and complexity of the hierarchical tables.

Factor analysis

The similarities between the groups of survey types (organization, ecosystem, substrate) were assessed with multiple correspondence analysis (MCA), using FactoMineR [28] and FactoShiny [29]. MCA is recommended for factor analysis of categorical variables instead of PCA [28]. We expected that the survey lists of similar types

(e.g., marine trash surveys) would use similar trash typology since they would have similar study goals. We split survey types by organization into research, nonprofit, and academic; ecosystems were split into marine, riverine, estuary, or land; substrate were split into beach, surface water, underwater, or roadside. First, we joined all classes used in each survey’s materials-items table to the alias tables. Second, we converted all classes to a matrix with zero denoting that the survey list did not have the class and one denoting that the survey list had the material and item class (one hot encoding). The MCA’s supplemental information (information not used to inform the model development) included organization, ecosystem, and substrate types (Table S3). V test statistics were assessed for each supplemental category’s first and second dimensions. V tests are used determine if a supplemental category has a MCA dimension significantly different from zero. We assigned a V test statistic value of 2 as the cutoff for significance (Table S4, Table S5).

Comparability analysis

We assessed the comparability of each survey list to all the others by calculating the one-way percent of

overlapping items or materials after joining them to the alias table:

tables (Fig. 3A, B). The item hierarchy was more complex than the material hierarchy. We have not yet developed

$$\text{Comparability Metric}_{X,Y} = \frac{\Sigma \text{Classes in sheet X equivalent with classes in sheet Y}}{\Sigma \text{All classes in sheet Y}} \quad (1)$$

where the *Comparability Metric*_{X,Y} is a one-way test for how comparable survey X is with survey Y. The metric defines the proportion of the classes in survey list Y that are accounted for by the classes in survey list X after joining the lists to the alias table. The comparability metric helps describe how much one survey accounts for the classes in another survey, a typical operation when merging trash survey lists. We then averaged all comparability metrics for each survey by material and item independently and plotted them to identify the most comparable surveys and discuss strategies for creating a 100% comparable survey list.

Another way to compare trash surveys is to lump them together using the hierarchy. We used the hierarchy and alias tables to compare the Stormwater Monitoring Coalition (SMC) survey list with the NOAA survey list. First, we added randomly sampled trash counts (a standard trash survey method) between 1 and 10 for each trash typology. We joined both surveys to the alias table and then to the hierarchical tables. We used the `data.tree` [25] package in R to sum up the hierarchies to demonstrate how the two surveys are related based on the hierarchy.

Results and discussion

The state of trash typology

Relational table summaries

Merging the material, item, and brand groups to their alias tables can inform us about the level of detail and potential applications for each group in use today. The alias lists condensed 87 classes to 25 unique materials, 1,138 items to 416 unique item classes, and 3,740 unique brand classes to 1,239 unique manufacturers. It is apparent that “material” is the most generic class used, “item” is more discretized, and “brand” has an even higher degree of subdivision. For survey development applications, reducing class choices to prime words in the alias lists helps to make surveys clearer and data more consistent. If users reduced classes to these alias terms before machine learning classification (e.g., the TACO trash classification routine) [30], they would improve their classification by clearly differentiating object classes and reducing labeling time.

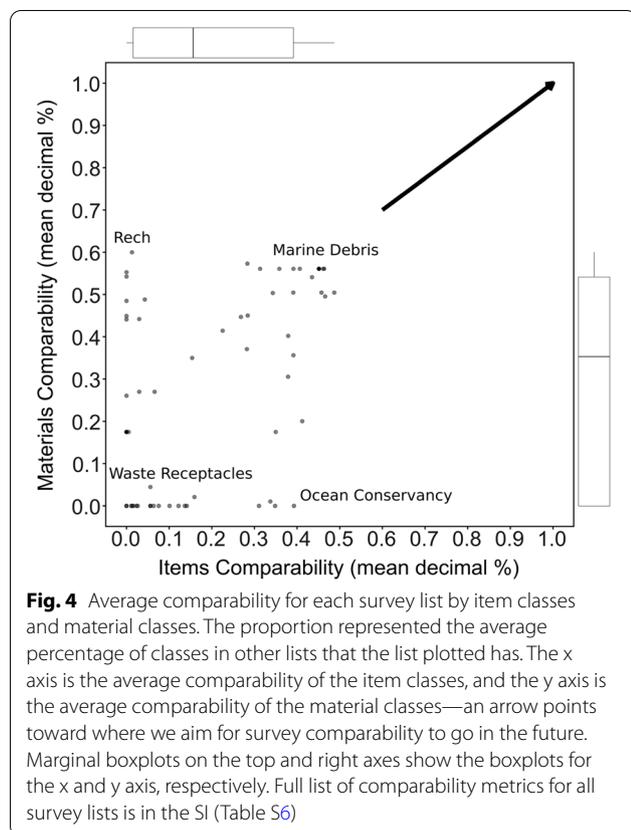
Inspecting the hierarchy tables can provide insight into the depth of information in the trash taxonomy and improve description clarity. There are four levels (parent–child word relationships) for material classes and six levels for item classes in the hierarchical relational

a hierarchy for brands, but we expect that one could be important for future developments. In an ideal hierarchical system, the terminal ends would encompass all possibilities of their higher class. For example, *fiberglass* would encompass all possibilities of *glass* (Fig. 3A). However, that is not the case here since we only characterized the surveys’ classes, and there are gaps in how trash surveys have characterized trash. Therefore, to accurately interpret the hierarchy, there is an implied *other* class as a subclass of each class wherever it is not explicitly made.

Out of 1509 material and item classes in the surveys, 392 did not fit our typology. We did not include them in our alias or hierarchical tables and instead made a separate misaligned class table. The main reason for misalignment was the categorization of trash by use, such as *fishing-related* or *construction materials*. A major challenge with these descriptors is that they can include a broad range of material, item, and/or brand classes, and thus do not fit within the framework we have developed. For example, the class *smoking-related* lumps item classes like *cigarette ends*, *tobacco packaging*, *matches*, *lighters*, and *pipes*, with material classes like *plastic*, *organic*, *paper*, *metal*, and *glass*. Additionally, while one organization may choose to include *lighters* as a *smoking-related* item, another may choose to put *lighters* in a *household item* use class. These descriptors could be useful for practitioners who want to conduct rapid assessments with lumped use-based descriptors, but we could not fold them into our tabular format because they defied the single parent rule for our hierarchical tables and the single alias rule for our alias tables. We recommend that future surveys designed to assess trash use class first start by describing the material and item classes and then build their use classes by summing observations of the material and item classes that fit their uses. These descriptors might be brought into alignment within the current system when a framework is developed that relates material–item–brand combinations to the use classes but would likely require a non-relational database schema due to the complexity of the relationships. The list of descriptors that did not fit the typology could be used to quickly filter less often used trash types and ambiguous trash typology during data mining routines.

Factor analysis of survey lists

Survey lists are often described as being for a specific type of organization, ecosystem, or substrate. We tested



whether those descriptors reflected differences in the suite of material class (46 survey lists) or item classes (52 survey lists) they describe using MCA (Table S3). No significant differences in material or item classes were found between survey lists by organization, ecosystem, or substrate type using the *v.test* statistic (Table S4, Table S5). This suggests that there was substantial overlap between the classes used in all types of surveys. In practice, a government marine survey list could be used for a nonprofit river survey, or any other combination of survey types, as long as it encompasses the project goals and is highly comparable with other survey lists.

Comparability analysis of survey list

In total, 4,556 comparability metrics were derived between all combinations of surveys for item and materials (Supplemental Information). The average comparability metric was 29.9% and 20.8% for all materials and items, respectively (Fig. 4 and Table S6). Some pairs of survey lists were 100% comparable for materials (562 pairs) and items (302 pairs). These lists could be compared directly without any inference or interpolation. A majority of the survey comparisons were 0% comparable for material (2,470 pairs) and item (2,104 pairs) classes. In these cases, list pairs are incompatible for data

aggregation and combination purposes at their current specificity level, a problem that cannot be rectified by the alias table alone. Only 15 of the 114 surveys were moderately comparable (> 40%) on average, and 33 surveys were somewhat comparable (40 – 20%) in item classes.

None of the surveys are 100% comparable with all others. However, the surveys produced by NOAA [31], SMC [30], OSPAR [32], Project AWARE [33], JRC [34, 35], Marine Conservation Society [36], and van der Velde [37] had the highest comparability values on average for material and item types. These surveys are potential candidates for adoption by new practitioners to enhance comparability. The JRC survey has the highest average for item classes (49% comparability), and the Rech [38] survey has the highest average for material classes (60% comparability). If the material classes from Rech [38] and the item classes from JRC survey were used in a survey, that survey would be more comparable than all others currently presented. The ultimate goal of harmonization is, in part, to achieve as much comparability between surveys as possible (Fig. 4).

The alias lists allow us to reach our goal of attaining more universally comparable surveys. If a survey contains all of the prime words for material and item classes, it will be 100% comparable to all of the surveys included in this study based on Eq. 1. An overlap in definitions between survey classes can be resolved by using the hierarchy as a part of the classification routine. This is done by choosing the most specific term within the hierarchical classification system (Fig. 1). This allows for more general and more specific terms to be used simultaneously. Users can already do this on the TTT website using the collapsible trees on the reference table tab. Figure 1 demonstrates how the hierarchy can be used to drill down to the most specific material and item classes that can describe the trash.

For those survey lists that are comparable, their typologies could be made even more similar by utilizing the hierarchy in addition to the alias to lump and split the classes. However, we do not see a clear “best” or “better” path forward within several possible options to lump and split datasets at the present moment. One strategy is to lump the values. All trash typologies will lump to the class *trash*, but in some cases, it is possible to lump to more specific classes (e.g., *forks* in one survey and *spoons* in another survey could lump to *utensils*) (Table 1). Another strategy to unify data is to split into more refined classes using the hierarchy tables, e.g., if one survey has *utensils* and the other survey has *forks*, *knives*, *spoons*, an attempt could be made to split the *utensils* class into more resolved classes. However, the analyst needs to have a way to infer what proportions the higher-level class should be split by to equal the values of the

Table 1 Example of lumping analysis for SMC material classes with the NOAA material classes. Both data cards are first joined to the materials alias relational table, then to the hierarchical table. The final result can be plotted to show how a joined survey list could be lumped. The tree structure under Hierarchical Classes is tabbed to show finer levels of granularity in the descriptions. Count is the number of objects of that material type that the surveys observed. Total Sum is how those sums would lump given the hierarchy. The class Missing is the number of classes that were not integrated into the TTT. Zero indicates an accurate merge

Hierarchical Classes	Count	Total Sum
• Litter	0	426
◦ Glass or Ceramic	0	30
■ Glass	30	30
• Metal	42	42
• Missing	0	0
• Organic	0	51
◦ Cloth	25	25
◦ Wood	26	26
• Other	85	85
• Plastic	223	254
◦ Rubber	31	31

refined class. This problem has yet to be solved for trash typology. Splitting may often not be possible because it requires additional information beyond the survey lists. A common challenge for lumping or splitting arises when a survey focuses on a particular set of items and materials but does not count the rest of the trash typology in an *other* class. The analyst then needs to infer the quantity and classes of trash that they did not characterize or only compare the quantities they did characterize to other studies. As a first step toward combining survey lists, we have attempted to solve the lumping problem using Table 1. This example demonstrates lumping counts from multiple surveys using different categories that are related by the hierarchy. We wrote an R script to do this automatically for any of the survey lists in the TTT (supplemental information). Although this demonstrates that survey lists can be merged by lumping programmatically, limitations due to previously stated method differences are likely to remain.

Conclusions

Future of the Trash Taxonomy Tool

Trash taxonomy will continue to evolve as new materials and items are created and enter environments [39] and as researchers create new technologies for collecting data about trash and develop new ways of describing trash [40, 41]. For further information on how the TTT will evolve with the addition of new survey data sheets,

as well as a current example with tobacco specific trash data, refer to the Practice Limitations section in the supplemental information. Our framework, relational tables, and source code will assist in developing and expanding the field of trash typology. Extensions to the TTT can be developed by directly collaborating on Github <https://github.com/wincowgerDEV/TrashTaxonomy> and submitting requests and feedback to <https://github.com/wincowgerDEV/TrashTaxonomy/issues>. Incorporating microplastic taxonomy is at the top of our list for taxonomy expansions [42]. The source code and data are licensed open access (CC by 4.0) attribution only. This analysis will need to be expanded to other languages in the future to accommodate differences in how different languages map the alias and hierarchy relationships. Translations are already being done with cross-country trash databases like the pan European Marine Litter Database [43]. Future work on database development should prioritize non-relational database structures, develop a reconciliation service in a standardized format [44], and assess the feasibility of incorporating semantic closeness and data value matching routines [27].

There is still much work to be done on the fundamentals of trash typology. Accurate brand identification is critical to ensuring the precise application of the principles of extended producer responsibility [45] to hold manufacturers accountable for large loads of post-consumption trash and substantial environmental impacts that result from the use of their products. We suspect that it will take an ongoing large-scale effort to keep up with the evolution of brand classes. Future work on brand classification within the TTT should include linking items, brands, and material combinations to identify the producers that ultimately should be responsible for the products they design and produce.

Several ongoing projects are using the TTT to assist in future developments and expand the current use cases (Table S7). The NOAA Marine Debris Program referenced the TTT in a recent update to their trash survey classes [46]. A recent study on roadside litter published a hierarchical categorization using the TTT to make their study results comparable to the other harmonized surveys [45]. The harmonized tables developed in this study are already being used to develop machine learning image classification in the Clean Currents Coalition so that the labels on trash items can be as restricted as possible without compromising the harmonizability of the dataset (personal communication). The Trash Monitoring Playbook [47] suggested using the TTT to trash survey practitioners.

The widespread adoption of the TTT would harmonize global efforts to measure and document trash loads, trash types, and the extent of trash pollution in environments.

The adoption of the TTT can also contribute to facilitating the aggregation of datasets from trash surveys, improving comparisons of trash risk assessments, and illuminating pathways for future work on trash taxonomy. We hope that TTT will be used to support research designed to inform mitigation efforts and prevention efforts, particularly in the realm of policymaking. We recommend using the TTT to foster collaborative research that will generate scientific evidence for holding producers accountable, ultimately by supporting “upstream” policy initiatives that reduce trash pollution of environments, promote changes in consumer behaviors, and mandate changes in producer practices.

Abbreviations

TTT: Trash Taxonomy Tool; JRC: Join Research Center; OSPAR: Oslo/Paris Convention; MCA: Multiple Correspondance Analysis; SMC: Stormwater Monitoring Coalition; NOAA: National Oceanic and Atmospheric Administration.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s43591-022-00035-1>.

Additional file 1. [48–50]

Acknowledgements

We want to thank the Trash Monitoring Work Group, which is part of the California Water Monitoring Council, for proposing these questions. We also thank the internal reviewers at NOAA who provided helpful feedback during their review, including Hillary Burgess, NOAA Marine Debris Program Monitoring Coordinator.

Authors' contributions

HH: Conceptualization, Methodology, Software, Validation, Formal Analysis, Investigation, Data Curation, Writing- Original Draft, Writing- Review & Editing, Visualization. WC: Conceptualization, Software, Methodology, Validation, Formal Analysis, Investigation, Data Curation, Writing- Original Draft, Writing- Review & Editing, Visualization. AG: Conceptualization, Validation, Writing- Original Draft, Writing- Review & Editing, Supervision. NT: Writing- Review & Editing. TH: Writing- Review & Editing. AM: Writing- Review & Editing. AV: Writing- Review & Editing. WY: Writing- Review & Editing. DA: Writing- Review & Editing. CM: Writing- Review & Editing. JV: Writing- Review & Editing. SS: Writing- Review & Editing. AH: Writing- Review & Editing. KY: Writing- Review & Editing. YK: Writing- Review & Editing. MM: Writing- Review & Editing. TL: Conceptualization, Software, Writing- Review & Editing. SM: Writing- Review & Editing. EB: Writing- Review & Editing. SL: Writing- Review & Editing, Funding Acquisition. PK: Writing- Review & Editing. GC: Writing- Review & Editing. JT: Writing- Review & Editing. JM: Conception, Investigation, Writing- Original Draft, Writing- Review & Editing. The author(s) read and approved the final manuscript.

Authors' Information

Not Applicable

Funding

This work was funded in part by a NOAA Marine Debris Program / National Marine Sanctuaries Foundation grant 19–09-B-233, USDA NIFA Hatch project number CA-R-ENS-5120-H, USDA Multi-State Project W4170, the Benioff Ocean Institute, and the UCR AES Mission Funding program. W. Cowger was funded by a National Science Foundation Graduate Research Fellowship and the McPike Zima Charitable Foundation. Dick Ayres was funded by the National Science Foundation. Jeremiah Mock was funded by the California Tobacco-related Disease Research Program Grant 28MT-0082. The scientific

results and conclusions, as well as any views or opinions expressed herein, are those of the authors and do not necessarily reflect the views of NOAA or the Department of Commerce.

Availability of data and materials

The tables and code are open access on Github to encourage users to update and improve on our relational framework (<https://github.com/wincowgerD/EV/TrashTaxonomy>). The tables and Shiny app tool have a CC BY 4.0 license. This gives users permission to remix, copy, and commercialize with attribution. Our intent is for others to build on these tools and use them widely.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent to publication

Not applicable.

Competing interests

The authors of this publication represent a diverse group of interests from regulators, regulatees, for profit companies, nonprofit organizations, government personnel, and academics who all have a stake in the problem of trash in the environment and want to see it ended. The corresponding author, Win Cowger, received no income from any invested entity. All authors confirm that none of these affiliations constitute actual conflicts of interest for the integrity of the manuscript. To see a full list of related funding for all coauthors please see below.

The Gray Lab (A. Gray, W. Cowger, and H. Hapich) would like to report that the funding they received from the Benioff Ocean Institute was 50% donated by the Coca Cola company, a plastic product producer, but the company had no input with the group about this publication. The Gray Lab has received funding through-flow from CalTrans, which is a regulated entity, but none of those funds were allocated to this project. The Gray Lab has received funding from Renew Oceans, a nonprofit that was supported by the Alliance to End Plastic Waste, a consortium of plastic producers, but the funds were not allocated to this project. The Gray Lab also receives funding from the NOAA Marine Debris Program, which is mandated to identify, determine sources of, assess, prevent, reduce, and remove marine debris. Win Cowger is on the Board of the Moore Institute of Plastic Pollution Research which is a nonprofit group focused on ending plastic pollution. Many authors are affiliated with the California Water Monitoring Council's Trash Monitoring Workgroup.

Margaret McCauley is the Trash Free Waters Program Lead for EPA's Region 10, the Pacific Northwest. In this role she endeavors to improve metrics for measuring and understanding what and how materials escape from management into waters of the U.S.

In recent years, Tony Hale and his colleague Shelly Moore served as co-PIs of the California Ocean Protection Council-funded Trash Monitoring Standards project. They also served as co-PIs on a project to devise a cigarette butt monitoring method, funded by the California Department of Public Health. The State Water Resources Control Board, an organization that regulates trash in waters, and the USEPA also fund the San Francisco Estuary Institute on a number of projects.

Neil Tangri is the Science and Policy Director at GAIA, a non-profit organization that advocates for stronger policies to reduce plastic pollution.

Litterati is a registered Delaware C Corp, the funding sources over the last several years have been from 3 sources 1. Private investment from accredited angel investors, 2. Grant funding from National Science Foundation under the SBIR grant and the Alliance to End Plastic Waste. 3. Commercial revenue from selling of product & services. None of these funding sources have supported or influenced the companies transparency goals or affected the input to this paper in determining the taxonomy mapping.

Sherry Lippiatt served as the NOAA Marine Debris Program (NOAA MDP) California Regional Coordinator, under contract with I.M. Systems Group and Lynker Technologies, during the development of this manuscript. The NOAA MDP, in partnership with the National Marine Sanctuary Foundation contributed funding to the development of the trash taxonomy tool.

Kathryn Youngblood is the Citizen Science Director for Marine Debris Tracker and a Research Engineer at the University of Georgia Circularity Informatics Lab. Debris Tracker (<https://debristracker.org/>) is a free mobile application

used by researchers, non-profit organizations, educators, and community members to collect geospatial data on marine, aquatic, and terrestrial debris. The app was developed with funding from NOAA in partnership with the University of Georgia and is now powered by Morgan Stanley. Data stewardship and operations are housed at the University of Georgia.

Gary Conley: Nothing to declare.

Yunfan Kang: Nothing to declare.

Antoinette Vermilye: Nothing to declare.

John Vermilye: Nothing to declare.

Charles Moore: Nothing to declare.

Jeremiah Mock: Nothing to declare.

Shelly Moore: Nothing to declare.

Walter Yu: Nothing to declare.

Eric Baggs: Nothing to declare.

Author details

¹Department of Environmental Science, University of California, 900 University Ave, Riverside, CA 92521, USA. ²Moore Institute for Plastic Pollution Research, 160 N. Marina Drive, Long Beach, CA 90803, USA. ³GAIA (Global Alliance for Incinerator Alternatives), Barangay Central Quezon City, Philippines.

⁴San Francisco Estuary Institute, 4911 Central Ave, Richmond, CA 94804, USA.

⁵University of California, 351 Winston Chung Hall, Riverside, CA, USA.

⁶Gallifrey Foundation, Chemin de l'Orchidee 2, 1299 CH Crans-pres-Celigny, Switzerland.

⁷Caltrans, 1120 N St, Sacramento, CA, USA. ⁸Litterati, 31 Turvey Ct, Chapel Hill, Chapel Hill, NC 27514, USA. ⁹Algalita, 160 N. Marina Drive, Long Beach, CA 90803, USA. ¹⁰Sacramento-San Joaquin Delta Conservancy, 1450 Halyard Dr., #6, West Sacramento, CA 95691, USA. ¹¹New Materials Institute, Center for Circular Materials Management, College of Engineering, University of Georgia, Riverbend Research Lab South, 220 Riverbend Road, Athens, GA 30602, USA. ¹²EPA, C/o U.S. EPA Region 10, 1200 Sixth Avenue, Suite 155, Seattle, WA 98101-3123, USA. ¹³University of California Riverside, 5809 Blackbird Ln, La Verne, CA 91750, USA. ¹⁴CA Waterboards, 11020 Sun Center Dr, Rancho Cordova, CA 95670, USA. ¹⁵NOAA Marine Debris Program / Lynker Technologies, 1301 Clay St, Suite 160N, Oakland, CA 94612, USA. ¹⁶CEFAS, Flat 6 Banner Court, Kirkley Cliff Road, Lowestoft NR33 0DB0, England. ¹⁷2nd Nature, 500 Seabright Ave, Santa Cruz, USA. ¹⁸Michael Baker International, 5 Hutton Centre Drive, Suite 500, Santa Ana, CA 92707, USA. ¹⁹University of California San Francisco, 490 Illinois St., Box 0646, San Francisco, CA 94143, USA.

Received: 8 October 2021 Accepted: 11 April 2022

Published online: 06 June 2022

References

- Laist DW. Impacts of Marine Debris: Entanglement of Marine Life in Marine Debris Including a Comprehensive List of Species with Entanglement and Ingestion Records. In: Coe JM, Rogers DB, editors. *Marine Debris: Sources, Impacts, and Solutions*. New York, NY: Springer New York; 1997. p. 99–139.
- Vegter AC, Barletta M, Beck C, Borrero J, Burton H, Campbell ML, et al. Global research priorities to mitigate plastic pollution impacts on marine wildlife. *Endanger Species Res*. 2014;25(3):225–47.
- Sigler M. The Effects of Plastic Pollution on Aquatic Wildlife: Current Situations and Future Solutions. *Water Air Soil Pollut Focus*. 2014;225(11):2184.
- Williams AT, Simmons SL. Sources of Riverine Litter: The River Taff, South Wales, {UK}. *Water Air Soil Pollut*. 1999;112(1):197–216.
- Hoellin T, Rojas M, Pink A, Gasior J, Kelly J. Anthropogenic litter in urban freshwater ecosystems: Distribution and microbial interactions. *PLoS One*. 2014;9(6):e98485.
- SFBRWQCB. California Regional Water Quality Control Board San Francisco Bay Region Municipal Regional Stormwater NPDES Permit, Order R2-2009-0074, NPDES Permit No. CAS612008. Adopted October 14, 2009. 279 pp. 2009;
- California Regional Water Quality Control Board Los Angeles Region. *Trash Total Maximum Daily Loads Los Angeles River Watershed*. 2007.
- Moore SM, Cover MR, Senter A. A Rapid Trash Assessment Method Applied to Waters of the San Francisco Bay Region: Trash Measurement in Streams. *SWAMP*. 2007.
- Prata JC, Silva ALP, Da Costa JP, Mouneyrac C, Walker TR, Duarte AC, et al. Solutions and integrated strategies for the control and mitigation of plastic and microplastic pollution. *Int J Environ Res Public Health*. 2019;16(13):2411.
- Kish RJ. Using legislation to reduce one-time plastic bag usage. *Econ Aff*. 2018;38(2):224–39.
- Lebreton L, Andrady A. Future scenarios of global plastic waste generation and disposal. *Palgrave Commun*. 2019;5(1):6.
- Allen K, Cohen D, Culver A, Cummins A, Curtis S, Eriksen M, Gordon M, Howe A, Lapis N, Prindiville M, Thorpe B, Wilson S. *Better Alternatives Now, {BAN} List 2.0*. 2017.
- Liao YH, Juang JG. Real-Time UAV Trash Monitoring System. *Appl Sci*. 2022;12(4):1838.
- Ocean Protection Council. *California Ocean Litter Prevention Strategy: Addressing Marine Debris from Source to Sea*. 2018.
- Morales-Caselles C, Viejo J, Martí E, González-Fernández D, Pragnell-Raasch H, González-Gordillo JI, et al. An inshore–offshore sorting system revealed from global classification of ocean litter. *Nat Sustain*. 2021;4(6):484–93.
- Vriend P, Roebroek CTJ, van Emmerik T. Same but Different: A Framework to Design and Compare Riverbank Plastic Monitoring Strategies. *Front Water* [Internet]. 2020 Oct 21;2. Available from: <https://www.frontiersin.org/articles/https://doi.org/10.3389/frwa.2020.563791/full>
- Fleet D, Vlachogianni T, Hanke G. A Joint List of Litter Categories for Marine Macrolitter Monitoring. 2021.
- Chang W, Cheng J, Allaire JJ, Xie Y, McPherson J. shiny: Web Application Framework for R [Internet]. 2020. Available from: <https://cran.r-project.org/package=shiny>
- Wickham H, François R, Henry L, Müller K. dplyr: A Grammar of Data Manipulation [Internet]. 2020. Available from: <https://cran.r-project.org/package=dplyr>
- Dowle M, Srinivasan A. data.table: Extension of `data.frame` [Internet]. 2020. Available from: <https://cran.r-project.org/package=data.table>
- Attali D. shinyjs: Easily Improve the User Experience of Your Shiny Apps in Seconds [Internet]. 2020. Available from: <https://cran.r-project.org/package=shinyjs>
- Chang W. shinythemes: Themes for Shiny [Internet]. 2018. Available from: <https://cran.r-project.org/package=shinythemes>
- Xie Y, Cheng J, Tan X. DT: A Wrapper of the JavaScript Library "DataTables" [Internet]. 2020. Available from: <https://cran.r-project.org/package=DT>
- Mason-Thom C. shinyhelper: Easily Add Markdown Help Files to "shiny" App Elements [Internet]. 2019. Available from: <https://cran.r-project.org/package=shinyhelper>
- Glur C. data.tree: General Purpose Hierarchical Data Structure [Internet]. 2020. Available from: <https://cran.r-project.org/package=data.tree>
- Khan A. collapsibleTree: Interactive Collapsible Tree Diagrams using "D3.js" [Internet]. 2018. Available from: <https://cran.r-project.org/package=collapsibleTree>
- Bellahsene Z, Bonifati A, Duchateau F, Velegarakis Y. On Evaluating Schema Matching and Mapping. In: Bellahsene Z, Bonifati A, Rahm E, editors. *Schema Matching and Mapping*. Berlin, Heidelberg: Springer Berlin Heidelberg; 2011. p. 253–91.
- Lê S, Josse J, Husson F. {FactoMineR}: A Package for Multivariate Analysis. *J Stat Softw*. 2008;25(1):1–18.
- Vaissie P, Monge A, Husson F. Factoshiny: Perform Factorial Analysis from "FactoMineR" with a Shiny Application [Internet]. 2021. Available from: <https://cran.r-project.org/package=Factoshiny>
- Moore S, Sutula M, Von Bitner T, Lattin G, Schiff K. *Southern California Bight 20 Regional Monitoring Program: Volume {III}. Trash and Marine Debris*. 2016.
- Opfer S, Arthur C, Lippiatt, Sherry, et al. *NOAA Marine Debris Shoreline Survey Field Guide*. NOAA Marine Debris Program. 2012.
- OSPAR. *Guideline for monitoring marine litter on the beaches in the OSPAR Maritime Area*. OSPAR Comm. 2010;1:84.
- Date S, Mm DD. *Dive Against Debris Data Card*. Project Aware. 2015.
- European Commission. *Institute for Environment and Sustainability. and MSFD Technical Subgroup on Marine Litter. Guidance for the Monitoring of Marine Litter in European seas*. [Internet]. 2013. Available from: <https://data.europa.eu/doi/https://doi.org/10.2788/99475>
- Hanke G, Galgani F, Werner S, Oosterbaan L, Nilsson P, Fleet D, Liebbezeit G. *Guidance on Monitoring of Marine Litter in European Seas: a guidance document within the Common Implementation Strategy for the Marine Strategy Framework Directive*. 2013.

36. Marine Conservation Society. Where does it all come from? [Internet]. 2012. p. 884. Available from: <https://ars.els-cdn.com/content/image/1-s2.0-S0048969716325918-gr11.jpg>
37. van der Velde T, Milton DA, Lawson TJ, Wilcox C, Lansdell M, Davis G, et al. Comparison of marine debris data collected by researchers and citizen scientists: Is citizen science data worth the effort? *Biol Conserv.* 2017;208:127–38.
38. Rech S, Macaya-Caquilpán V, Pantoja JF, Rivadeneira MM, Jofre Madariaga D, Thiel M. Rivers as a source of marine litter - A study from the SE Pacific. *Mar Pollut Bull.* 2014;82(1–2):66–75.
39. Prata JC, Silva ALP, Walker TR, Duarte AC, Rocha-Santos T. COVID-19 pandemic repercussions on the use and management of plastics. *Environ Sci Technol.* 2020;54(13):7760–5.
40. Zhao Z, Kang Y, Magdy A, Cowger W, Gray A. A Data-Driven Approach for Tracking Human Litter in Modern Cities. In: 2019 IEEE 35th International Conference on Data Engineering Workshops (ICDEW). IEEE; 2019. p. 69–73.
41. Kang Y, Zhao Z, Magdy A, Cowger W, Gray A. Scalable multi-resolution spatial visualization for anthropogenic litter data. In: GIS: Proceedings of the ACM International Symposium on Advances in Geographic Information Systems. 2019.
42. Helm PA. Improving microplastics source apportionment: a role for microplastic morphology and taxonomy? *Anal Methods.* 2017;9(9):1328–31.
43. Molina Jack ME, Chaves Montero M del M, Galgani F, Giorgetti A, Vinci M, Le Moigne M, et al. EMODnet marine litter data management at pan-European scale. *Ocean Coast Manag.* 2019;181:104930.
44. Delpeuch A, Pohl A, Steeg F, Sr. TG. Reconciliation Service (API) [Internet]. Available from: <https://reconciliation-api.github.io/specs/0.1/>
45. Cowger W, Gray A, Hapich H, Osei-Enin J, Olguin S, Huynh B, et al. Litter origins, accumulation rates, and hierarchical composition on urban roadsides of the Inland Empire, California. *Environ Res Lett* [Internet]. 2022;17(1):15007. Available from: <https://doi.org/10.1088/1748-9326/ac3c6a>.
46. Burgess HK, , Charlie E. Herring, Sherry Lippiatt, Sarah Lowe and AVU. NOAA Marine Debris Monitoring and Assessment Project Shoreline Survey Guide. 2021;(April):1–20.
47. Moore S, Hale T, Weisberg SB, Flores L, Kauhanen P. California Trash Monitoring Methods and Assessments Playbook. 2020.
48. Mock J, Hendlin YH. Notes from the Field: Environmental Contamination from E-cigarette, Cigarette, Cigar, and Cannabis Products at 12 High Schools - San Francisco Bay Area, 2018–2019. *MMWR Morb Mortal Wkly Rep.* 2019;68(40):897–9.
49. Lee YO, Glantz SA. Menthol: putting the pieces together. *Tob Control* [Internet]. 2011 May 1;20(Supplement 2):ii1–7. Available from: <https://tobaccocontrol.bmj.com/lookup/doi/https://doi.org/10.1136/tc.2011.043604>
50. Pirika. Towards Litter-Free Cities [Internet]. Available from: <https://en.research.pirika.org/>

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ [springeropen.com](https://www.springeropen.com)
