



Article

Life Cycle Data Interoperability Improvements through Implementation of the Federal LCA Commons Elementary Flow List

Ashley N. Edelen ^{1,*}, Sarah Cashman ¹ , Ben Young ¹  and Wesley W. Ingwersen ² ¹ Eastern Research Group, Inc., Lexington, MA 02421, USA² Center for Environmental Solutions and Emergency Response, Office of Research and Development, United States Environmental Protection Agency, Cincinnati, OH 45220, USA

* Correspondence: ashley.edelen@erg.com

Abstract: As a fundamental component of data for life cycle assessment models, elementary flows have been demonstrated to be a key requirement of life cycle assessment data interoperability. However, existing elementary flow lists have been found to lack sufficient structure to enable improved interoperability between life cycle data sources. The Federal Life Cycle Assessment Commons Elementary Flow List provides a novel framework and structure for elementary flows, but the actual improvement this list provides to the interoperability of life cycle data has not been tested. The interoperability of ten elementary flow lists, two life cycle assessment databases, three life cycle impact assessment methods, and five life cycle assessment software sources is assessed with and without use of the Federal Life Cycle Assessment Commons Elementary Flow List as an intermediary in flow mapping. This analysis showed that only 25% of comparisons between these sources resulted in greater than 50% of flows being capable of automatic name-to-name matching between lists. This indicates that there is a low level of interoperability when using sources with their original elementary flow nomenclature, and elementary flow mapping is required to use these sources in combination. The mapping capabilities of the Federal Life Cycle Assessment Commons Elementary Flow List to sources were reviewed and revealed a notable increase in name-to-name matches. Overall, this novel framework is found to increase life cycle data source interoperability.

Keywords: nomenclature; life cycle assessment; elementary flow; interoperability



Citation: Edelen, A.N.; Cashman, S.; Young, B.; Ingwersen, W.W. Life Cycle Data Interoperability Improvements through Implementation of the Federal LCA Commons Elementary Flow List. *Appl. Sci.* **2022**, *12*, 9687. <https://doi.org/10.3390/app12199687>

Academic Editors: Arcangelo Castiglione and Dimitris Mourtzis

Received: 18 July 2022

Accepted: 21 September 2022

Published: 27 September 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Increasingly, life cycle assessment (LCA) is being used as an environmental management tool to make key decisions in reducing environmental impacts and optimizing product systems. LCA as a decision support tool is very data intensive, and most models utilize data from multiple data sources to provide the necessary background data or life cycle inventory (LCI) data to accompany the primary data used to model a product's life cycle [1]. The lack of harmonization across different data providers, software providers and life cycle impact assessment (LCIA) methods is a well-documented challenge [2–8]. Previous research efforts have been investigated to highlight and showcase the myriad of challenges associated with LCA data interoperability [9], even going so far as to create methodologies to review and assess the level of harmonization across various platforms [9,10]. Past research has highlighted the issues around nomenclature [7,10] and the need for improved semantics within LCA to access machine capabilities for harmonizing data across platforms [4,9,10]. These improvements are particularly important for elementary flows (EF), which are the foundational component of LCAs and link LCI and LCIA data to generate results.

Efforts on a global and U.S. national scale to harmonize LCA data have been ongoing since the early 2000s [11]. Global efforts for the harmonization of data in LCA are driven by the Life Cycle Initiative of the United Nations Environment Programme (UNEP) through

the Global Life Cycle Data Access (GLAD) initiative [12]. This initiative supported a global effort to analyze interoperability issues across LCA software, LCI databases and LCIA methods, and published the findings of these analyses in a critical review published in 2018 that produced 11 recommendations [10]. The evaluation of EF interoperability has been further improved with the methodology to develop pairwise mapping between major elementary flow lists developed through a project of the GLAD nomenclature working group [13].

On a national level, U.S. federal agencies have worked collaboratively through the Technical Working Group (TWG) on Federal LCA Data Interoperability initiated in 2014 to harmonize data across U.S. national agencies to produce a common data portal called the Federal LCA Commons [14–16]. As a part of this effort, the Federal LCA Commons decided to create a harmonized nomenclature of EFs called the Federal Elementary Flow List (FEDEFL). The objective of this paper is to analyze the ability of the adoption of the novel FEDEFL, built from the critical review recommendations [10], to effectively improve interoperability for LCA software, databases and LCIA method providers. The authors provide key elements of the FEDEFL framework in Section 2 to highlight the key components of the FEDEFL that were part of the analysis.

2. State of the Art of the Federal Elementary Flow List

This section highlights key components of the FEDEFL framework that are essential to the analysis. Detailed information on the FEDEFL structure can be found in a U.S. Environmental Protection Agency (USEPA) report [1]. The FEDEFL consists of a component-based structure, organizing and separating different data and metadata. The importance of this structure is to allow for increased automation in the handling of the list.

2.1. Federal Elementary Flow List Components

The creation of the FEDEFL was motivated by the need to be able to combine the individual components of a flow and eliminate the inconsistent inclusion of information in flows. There are two main types of components, the components of a flow and the metadata components, which are defined below.

2.1.1. Elementary Flow Components

As defined by the FEDEFL USEPA report [1], EFs must have three components to identify them, which are required for a flow to appear in the FEDEFL:

1. Flowable—The name of the material, energy, or space (e.g., “Carbon dioxide” or “freshwater”) that comes from or goes to the biosphere.
2. Context—A set of environmental media/compartments that describe the flow origin or destination.
3. Unit—Flow units may be associated with conversion factors that can be used to convert between different units within a flow property (e.g., kg to lbs.) or even between flow properties (e.g., kg to m³).

2.1.2. Flow Metadata

Flow metadata consist of several components, including clarifiers, identifiers, unit converters, and secondary context information. Table 1 defines and provides examples of each component.

Table 1. Components of flow metadata.

Metadata Component	Definition	Example(s)
Clarifiers		
CAS Number	A unique numerical identifier assigned by the Chemical Abstracts Service	10102-44-0
Formula	Identifies each constituent element by its chemical symbol and indicates the proportionate number of atoms of each element	NO ₂
Synonyms	Another name for the same flowable	Nitrogen oxide
External Reference	External definition	Information on the substance details from USEPA's Substance Registry Services (SRS)
Identifiers		
Flow UUID	A universally unique identifier to identify a flow (flowable + context)	b285ccbd-7703-39d9-9f66-20451f18d99f
Unit Converter		
Alternate Unit	A unit other than the default unit for a flowable	kg
Conversion Factor	An arithmetical multiplier for converting from default to alternative units	37 kg/MJ

Defining the structure and the inclusion of metadata with flowables helps to eliminate metadata being mixed within the flow names. When EF lists mix metadata and flowables, the names become complex and automatic matching between flowables decreases significantly. This simplifying of the structure by defining the metadata components as separate from the flowable allows for increased automatic matching.

2.2. Flow List Structure

Flow Classes

The GLAD critical review developed flow classes as an additional flow metadata component, and these flow classes are used to organize flows and create common naming structures by flow class [10]. These flow classes were further refined in the FEDEFL USEPA report [1]. Flow classes are a way to group EFs by their potential uses in LCA data. Classes may have sets of contexts and units that distinguish them from flowables in other classes. Table 2 is the modified version of the flow classes used in the FEDEFL USEPA report [1].

Table 2. Flow classes.

Flow Class	Resource/Emission	Definition	Example Flowable(s)
Biological	Both	Biomass or organic matter (i.e., microorganisms)	'Wood' or 'Bacillus subtilis'
Element or Compound	Both	A unique chemical element or compound	1,1,2,2-Tetrachloroethane
Energy	Both	Energy input NOT associated with consumed materials including heat	Energy, Geothermal
Geological	Both	A mineral or metal in an ore or aggregate material extracted for use or refining	Anthracite
Groups of Chemicals	Both	A group or mixture of chemicals	Dioxins
Land Use	Input	Land types	Land
Water	Both	Water	Water, fresh
Other	Both	None of the above. May include water quality parameters	Biological oxygen demand

Flow classification is an extremely useful tool during flow mapping between LCA data sources, especially for large EF lists. By classifying the lists into smaller sub-groups, it is significantly easier to complete manual matching that may be necessary during a mapping process.

2.3. Source Data

To improve source interoperability, a systematic approach was taken to collect flowable names from federal emission, LCA and LCIA sources. Sources were expanded to make the list interoperable with other U.S. government-supported data sources. Including multiple types of sources in the creation of the FEDEFL improves the coverage of the EF list. This robust method of systematically building an EF list is seen as improving the ability of the FEDEFL to operate as a common source flow list. The following types of sources and sources were used in the creation of the FEDEFL.

2.3.1. USEPA Sources

- TRI—Toxic Release Inventory. The TRI is a USEPA-produced data source for tracking potential toxic wastes and releases which includes available emission data for industrial facilities that are either from a specific industry, employ 10 or more full-time employees, or manufacture, process or otherwise use TRI-listed chemicals [17].
- NEI—National Emissions Inventory. The NEI is an estimate of criteria pollutants, criteria precursors, and hazardous air pollutants built using the Emissions Inventory System to collect and blend data from state, local, and tribal air agencies, and is updated every three years [18].
- RCRA—Resource Conservation and Recovery Act. RCRA is a USEPA reporting program to track non-hazardous solid waste and hazardous solid waste from ‘cradle to grave’ that requires large-quantity generators to report every two years [19].
- DMR—Discharge Monitoring Report. The National Pollution Discharge Elimination System (NPDES) collects reports for point source discharges to water bodies from permitted facilities [20]. DMR is a collection of the periodic (monthly, seasonally or semi-annual) water pollution reports derived from NPDES.
- eGRID—Emissions & Generation Resource Integrated Database. eGRID is a data source for air emissions and the generation of electrical power in the U.S. [21].
- PDP—the Pesticide Data Program. The USDA supports agricultural programs such as the PDP which monitors the residues of pesticides on agricultural products [22].
- Mineral Commodities Summary. The Mineral Commodities Summary is an annual report published by the U.S. Geological Survey covering the non-fuel mineral industry [23].
- Water Data for the Nation. The Water Data for the Nation provides water resource data about occurrence, quantity, quality, distribution, and movement of surface and underground waters [24].

2.3.2. Database Sources

- GREET—The Greenhouse gases, Regulated Emissions, and Energy use in Technologies (GREET) Model created by Argonne National Laboratory [25]. GREET is a full LCA model supporting well to wheel, and fuel and vehicle cycles through material and disposal, covering energy and emissions for advanced and new transportation fuels.
- USLCI—U.S. Life Cycle Inventory (U.S. LCI) database by the National Renewable Energy Laboratory [26]. The U.S. LCI database is a collection of individual gate-to-gate, cradle-to-gate and cradle-to-grave life cycle assessments supporting material, component, or assembly within the U.S. The U.S. LCI is a repository for different North American stakeholders (e.g., industry associations) to provide public LCI data.
- Electricity Baseline LCI. The Federal LCA Commons Baseline [27] U.S. Regional Electricity LCI data available at <https://www.lcacommons.gov/> (accessed on 4 June 2021).

2.3.3. Software Sources

- openLCA—openLCA is a free open-source life cycle assessment software [28]. The openLCA EF list was utilized as an original source and the software has been used as a tool to support mappings. Training material has been developed to describe mapping to different EF lists, including the FEDEFL, within openLCA [29].

2.3.4. Life Cycle Impact Assessment Method Sources

The FEDEFL operates with common LCIA methods. The companion USEPA LCIA formatter tool (<https://github.com/USEPA/LCIAformatter> (accessed on 13 September 2021)) was designed to use the FEDEFL as the EF structure for all LCIA and LCI methods processed through this package [30]. The FEDEFL and LCIA formatter process the following LCIA methods:

- TRACI—Tool for Reduction and Assessment of Chemicals and other environmental Impacts (TRACI). TRACI is a USEPA developed environmental impact assessment tool for characterization of life cycle data [31].
- ReCiPe—An impact assessment method for LCA first developed in 2008 through a joint effort from RIVM, Radboud University Nijmegen, Leiden University and PRÉ Sustainability [32].
- ImpactWorld+—A globally regionalized LCIA method based on a midpoint damage framework with four distinct complementary viewpoints [33].

2.3.5. Capturing Source Flows

The FEDEFL was designed to be an all-encompassing list capable of handling all flows to and from nature. The automated Standardized Emissions and Waste Inventories (StEWI) module developed by USEPA was used to generate comprehensive lists of flows used in USEPA datasets (i.e., inventory sources described above) [34]. Once chemicals were collected from these sources, chemicals were defined using two USEPA chemical databases, the Substance Registry Services (SRS) [35] and the Chemistry Dashboard [36]. SRS is the USEPA's authoritative resource on chemicals, biological organisms and other substances tracked and regulated by USEPA. The Chemistry Dashboard is a database for chemistry, toxicity, and exposure information for over 760,000 chemicals. The USEPA chemical databases were used to match chemical names, CAS No., and chemical formulas. SRS names are used in preference over chemistry dashboard naming. A common naming system for flows allows for the removal of duplicates and the correspondence of many sources of flowables and contexts to the FEDEFL.

2.3.6. Available Mappings

Mapping files are the essential tools of harmonization across platforms. Mapping files connect one source to another, usually at a flow level. Flow mappings are provided within FEDEFL software package for several commonly used existing LCI or LCIA sources to ensure compatibility with the FEDEFL and each other. In many cases, mapping files are generated by matching flowables and contexts separately. This reduces the potential of inconsistent matching across flowables when tracked in multiple contexts and inconsistent contexts. Upon generating a mapping file, flow Universal Unique Identifiers (UUID) are checked to ensure that the target flowable and context exists in the FEDEFL. The methods section describes the simplistic mapping method used to determine interoperability for the analysis.

There are currently several available mapping files. Existing mapping files were used to support the analysis for this paper and guide the selection of the sources for the analysis. Currently, mapping between the FEDEFL and the GREET model, USLCI database version 2019Q4, and TRACI 2.1 are available. The open-source software openLCA has been working with the Federal LCA Commons to integrate flow list management changes; therefore, a mapping to the reference list from openLCA is also available [29]. Since openLCA is a repository for various datasets, original data sources connected with the openLCA reference list were also not a part of the mapping process. In addition to Federal LCA Commons sources, mappings to LCIA methods ReCiPe and ImpactWorld+, and LCA sources IDEA and ecoinvent, are available. All these mappings are one-way mappings to the FEDEFL and are available online on the FEDEFL github repository. One-way mappings are mappings that have been completed as a part of the Federal LCA Commons and have not in the

instances of ReCiPe, ImpactWorld+, IDEA or ecoinvent been reviewed or approved by the original source.

3. Materials and Methods

Three analyses were conducted on the FEDEFLL version 1.0.8 to test interoperability improvements. The analyses employed the same methods to match EFs or components of these flows as described in previous studies [10,13]. The analyses used publicly available sources that were available to the authors. The first analysis was conducted to determine how capable the FEDEFLL is at mapping to other LCA sources. Two databases and datasets (ecoinvent and USLCI), one software (openLCA), and two LCIA methods (TRACI and ImpactWorld+) with existing mappings to FEDEFLL were selected for analysis. The number of unique flowables that were mapped were compared with the total number of unique flowables in each of the sources. In this analysis, unique flowables refers to the name of the flows, excluding any context or unit information.

The second analysis was performed to determine the current interoperability between other LCA EF lists. The sources of EF lists is limited to either publicly available EF lists or EF lists that were available through previous and ongoing collaborative efforts, such as the critical review [10,13]. The authors acknowledge that additional databases and datasets, software and LCIA methods exist, but were not included in the comparison. This comparison is not meant to be an exhaustive comparison of all LCA sources, but a representation of general issues between LCA sources. Flows from each list were compared to flows in every other flow list using name-to-name matching. For the purposes of simplicity, context information was not included when matching flows, and only unique flows (i.e., flowables) were used from each list. Ten sources were selected for this analysis: four databases and datasets, three pieces of software and three LCIA methods. Table 3 provides the ten source names and versions that were used for this analysis

Table 3. Flow list analysis sources.

Source Type	Source Name	Version
Databases and datasets	USEEIO	v1.1 [37]
Databases and datasets	USLCI	Q4v2019 [26]
Databases and datasets	IDEA	v2.3 [38]
Databases and datasets	ecoinvent	v3.6 [39]
Software	openLCA	Reference flow list (2020) [28]
Software	SimaPro	Used in 2018 GLAD critical review [10]
Software	GaBi	Used in 2018 GLAD critical review [10]
LCIA methods	TRACI	v2.1 [31]
LCIA methods	ReCiPe	2016 release [32]
LCIA methods	ImpactWorld+	[33]

Lastly, an analysis was conducted to use the FEDEFLL as an intermediate flow list in mapping LCI databases to LCIA methods. Flow lists from two LCI databases—USLCI and ecoinvent—and lists from three LCIA methods were used in the analysis. The versions of the LCI and LCIA sources used are listed in Table 3. For this analysis, the source lists were both mapped to FEDEFLL and then compared to see if there was an improvement in matching capabilities between the original source to source matching and the matching using the FEDEFLL as a common EF list.

4. Results

Table 4 shows how many flows from each of the five sources were capable of being mapped to the FEDEFL from the total flows within the source. Mappings were possible for 85–98% of flows from these five other flow lists. Most of the flows that remained unmappable either contained information that was not possible to capture with the FEDEFL, such as long-term emissions in ecoinvent, or flows that were unidentifiable, such as chemicals that lacked CAS numbers.

Table 4. Available mappings analysis.

	Ecoinvent	USLCI	OpenLCA	TRACI	ImpactWorld+
Mapped	3751	4351	4605	3568	3457
Total	4323	4450	5216	3950	4048
% of flows mapped	87%	98%	88%	90%	85%

In Table 5, an analysis comparing flow names from different sources shows some of the challenges associated with interoperability. The table in Table 5 is read as the percentage of flowables found in the source list in the first column that can be found in the source list in the top row. For example, 55% of flowables from ecoinvent can be found in USLCI. Lists were viewed as having significant matching capabilities if greater than 50% of unique flowables were found to match. Lists were defined as having less than marginal matching capabilities if less than 25% of unique flowables matched. Based on name–name matching between sources, 25% of dataset comparisons had significant matches, while 56% of comparisons showed less than marginal matches.

An example of significant flow matching capabilities can be found for USEEIO flowables found in openLCA, as 64% of flowables from USEEIO matched with a flowable in openLCA. IDEA was found to have marginal matching capabilities with openLCA with 31% of flowables matching. In fact, nine sources found marginal matching capabilities with openLCA and seven out of nine sources found significant matching capabilities with openLCA. openLCA had the greatest capacity for automated name–name matching (defined as the greatest number of significant and marginal matching capabilities based on openLCA column). It is to be expected that when other sources are matched with software sources, higher matching capabilities will be observed. This is because software EF lists tend to be created as an amalgamation of EFs from other sources and are more inclusive in their structure.

Table 5. Comparison of interoperability between LCA sources.

Legend		Databases and Datasets				Software			LCIA Methods		
	<25%	USEEIO ¹	USLCI ²	IDEA ³	Ecoinvent ⁴	OpenLCA	SimaPro ⁵	GaBi ⁵	TRACI ⁶	ReCiPe ⁷	ImpactWorld
Databases and Datasets	USEEIO ¹		30%	14%	34%	64%	51%	11%	44%	41%	44%
	USLCI ²	23%		7%	18%	74%	54%	17%	53%	18%	32%
	IDEA ³	27%	17%		16%	31%	27%	14%	25%	26%	25%
	Ecoinvent ⁴	5%	55%	1%		69%	11%	2%	26%	25%	39%
Software	OpenLCA	9%	13%	2%	12%		59%	4%	32%	28%	42%
	SimaPro ⁵	8%	11%	2%	12%	66%		3%	34%	30%	44%
	GaBi ⁵	19%	37%	13%	25%	45%	34%		17%	19%	27%
LCIA Methods	TRACI ⁶	6%	7%	2%	10%	71%	30%	1%		75%	63%
	ReCiPe ⁷	15%	9%	6%	12%	73%	69%	5%	86%		71%
	ImpactWorld	12%	13%	4%	16%	95%	79%	4%	64%	62%	

¹ USEEIOv1.1; ² USLCIQ4v2019; ³ IDEA 2.3; ⁴ ecoinvent 3.6; ⁵ SimaPro and GaBi from the 2018 GLAD critical review; ⁶ TRACI 2.1; ⁷ ReCiPe 2016.

Databases and datasets tend to be narrower in scope and are, therefore, most likely to exhibit the lowest matching rates, as seen in the first four columns of Table 5 (under the databases and datasets header). However, when databases and datasets are matched with software sources, the larger, more inclusive nomenclatures of the software tend to allow for higher matching rates. The low number of significant matches found between databases and datasets and LCIA methods (found in the last three columns of the table) shows why it is important for mapping files to exist between sources. A lack of connectivity between databases, datasets, and LCIA methods can lead to inaccurate LCA results. Therefore, most software sources have a mapping file with LCIA methods. However, mappings can be time consuming to produce and need to be consistently updated for interoperability between sources to be maintained.

When the FEDEFLL is used as an intermediate mapping, more flows from USLCI andecoinvent Figure 1 map to LCIA methods. The matching percentage between LCI databases and LCIA methods increased 38% to 277% across the different LCI and LCIA sources.

The variation in improvement in matching can be attributed in large part to the framework and structure of the original sources. Some of the databases and LCIA methods have similar naming structures or nomenclature rules when compared with other sources. This could account for minimal increases in interoperability in select cases when using a common standard list. Additionally, it is expected that not all flows should match between an LCI source and LCIA method, as there are some flows that will not result in an impact for the categories assessed in the selected LCIA method. By adhering to the strict definition of an EF used by the FEDEFLL, not all the original source EFs in USLCI are considered EFs. For example, 'volume occupied, reservoir' or 'process effluent' from the USLCI are not mapped to the FEDEFLL since the FEDEFLL does not consider space occupation an EF and process effluent is also not a specific substance, but rather represents a mix of unknown substances. In the case of 'volume occupied, reservoir', this also does not correspond to impact assessment categories in the covered LCIA methods, so exclusion is not expected to affect the robustness of study results. In the case of 'process effluent,' the original LCI developer did not provide sufficient information on the composition of the flow to allow for incorporation in an LCIA method. While tools such as the FEDEFLL improve the interoperability and structure of LCA models, model results remain dependent on the quality of the underlying source data.

As the GLAD critical review and Table 5 revealed, there currently exists a gap between LCA sources' EF structures and LCIA methods' EF development [10]. The use of a standard list, such as the FEDEFLL can be a vital step forward to closing the gap between LCIA method development and LCI databases.

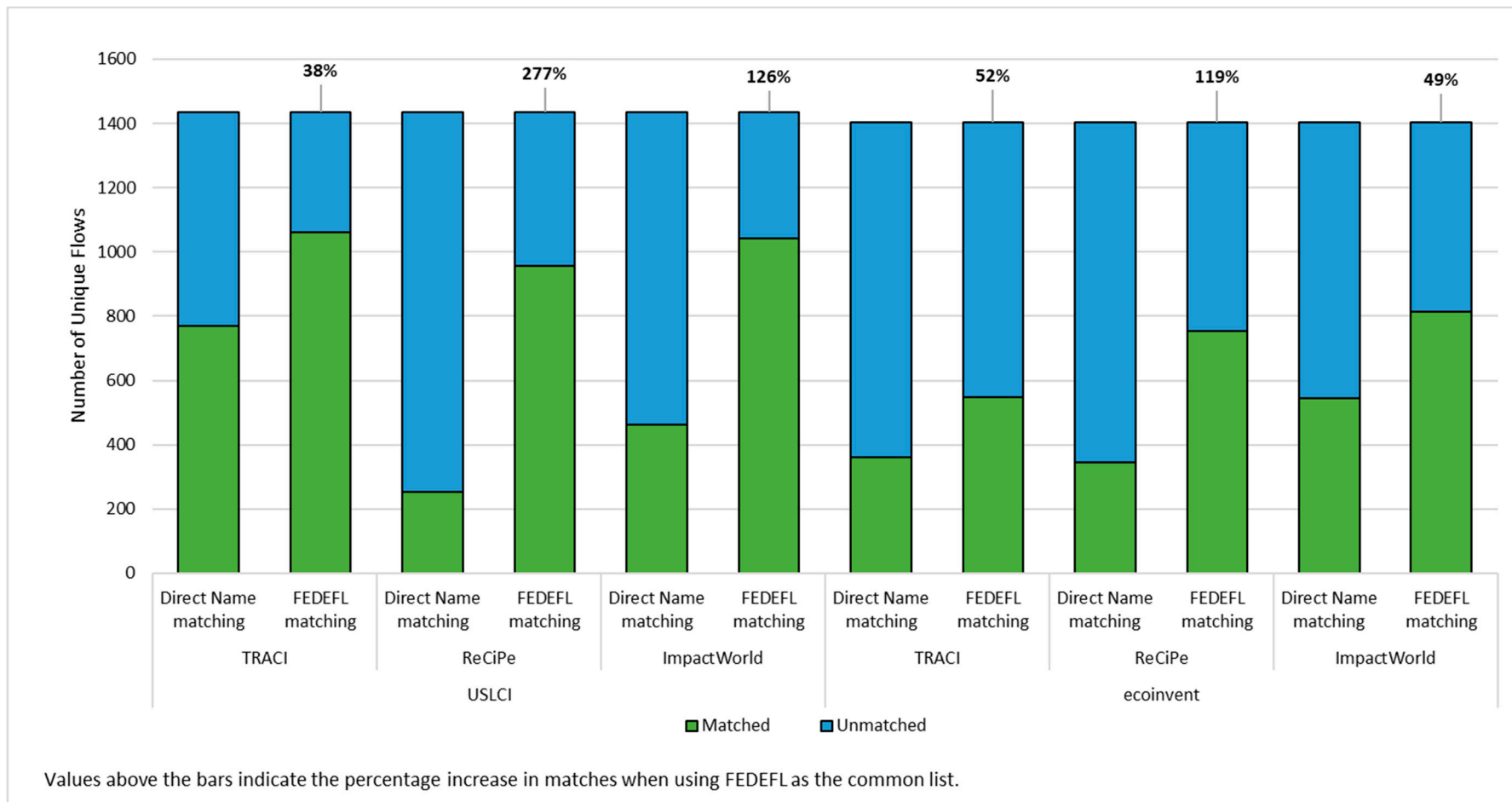


Figure 1. Analysis of improved matching using FEDEFL as the common list.

5. Discussion

Due to the diverse nomenclatures and LCA development methodologies that are currently available, users face a significant challenge in synthesizing and harmonizing data from different sources within the LCA community. As shown in Table 5 significant differences in nomenclature between sources still exist, as found by Edelen et al. [10]. This lack of EF connectivity can lead to flows inadvertently being excluded from an LCIA analysis, which presents significant challenges to LCA developers and can lead to erroneous results. In the past, the only alternative for developers is to spend significant resources to manually map EFs for each individual LCA. With the automated tools in the FEDEFLL toolbox, users can greatly decrease the resources necessary to harmonize EFs and improve the accuracy of LCA study results.

Users should also be aware of the limitations of mapping files. The mapping files available through the FEDEFLL repository are one way mapping files; mappings are completed by FEDEFLL to map from other data sources to the FEDEFLL. While a flow may be mappable in one direction, it does not mean that the reverse mapping is accurate. In mapping files, flows are mapped using different mapping conditions, "=", "~", ">", and "<". While some flows are ideal matches and can be matched using the "=" signs, this is not always true for all mappings. For example, in one source you may see the flowable name "Silver(II)" and in another list they only have "Silver". The "Silver" flowable is not an exact match to "Silver(II)" and can include more silver ions that just "Silver(II)"; therefore, a "<" with the greater side directed at "Silver" would be used. This type of mapping is performed at the discretion of the mapping creators and often is performed in instances where the chemical has a known impact. However, not all flows are mapped to proxy flows and users should be careful when using mappings as some flows with known impacts may disappear during the mapping process or end up mapped to flows that the user does not find to be an acceptable proxy flow.

The mapping files posted through the FEDEFLL were created by developers of the FEDEFLL and not necessarily with the collaboration of the original data sources. This creates limitations as flows from the original sources may have been unclear or misinterpreted by the FEDEFLL developers due to lack of metadata. Examples of developer-dependent decisions include mapping flows with errors, such as misspelled flows. Often developers are left to either leave misspelled flows out of mappings or map to the flow they believe the original data source was referencing. Mapping developers must make decisions such as whether to map flows with greater specificity to flows with lesser specificity. This can occur for both context information and for flowable information. An example of this challenge is particulate matter. Some databases have flows such as "Particles (PM2.5)", "Particles (PM0.2–PM2.5)", and "Particles (PM0.2)" where the FEDEFLL only has "Particulate matter, $\leq 2.5 \mu\text{m}$ ". An example where context information may be more specific is the inclusion of some databases of flows with "long-term" context subcompartment and the exclusion from the FEDEFLL. Decisions such as these create significant challenges for projects such as that undertaken by the GLAD nomenclature working group to provide dual direction mapping files from four LCA developers (i.e.,ecoinvent, ILCD, FEDEFLL and IDEA), as different developers must spend significant time agreeing on these decisions. The GLAD nomenclature working group has provided a report of the challenges faced during this project in addition to the mapping files so that users may better understand the limitations of mappings.

The removal of exchange information from flow names does have an impact on the mapping abilities of FEDEFLL flows with other EF lists and ultimately creates challenges with LCA studies. The inclusion of exchange information in a flow name, such as 'biogenic' or the flow location, allows software to treat these flows as separate and, therefore, trace impacts separately. However, with the removal of this information these flows are no longer tracked within a LCA system and a LCA developer must manually track these flows to include in an impact assessment. Including such exchange information in a flow name can, however, lead to long and complicated flow lists. For example, many water flows

include location name to be able to track regional water scarcity. Future improvements of the different LCA software tools could instead link the location to the exchange rather than directly to the flow name. Currently, an alternative to manual tracking of these flows does not exist.

6. Conclusions

The future efficacy of LCA as a decision support tool requires investments to improve interoperability. However, there are limitations to interoperability if the LCA community continues to lack consensus on adopting a single nomenclature. Mapping files will have to be consistently produced and updated as source lists evolve, and new versions are released. Manual mapping of certain flows will continue to play a critical and time-consuming role in interoperability. A lack of consistent structure and nomenclature around flow context information also adds complexity to the interoperability of EF nomenclature.

The development of the FEDEFL framework and the resulting flow list was demonstrated to improve interoperability. The structure of the FEDEFL nomenclature has been used as a roadmap for other LCA databases and data developers to improve the interoperability of lists through the increased use of metadata to identify flows and through the utilization of flow components to simplify the flowable names. The novel FEDEFL framework is adaptable and has been utilized in the development of the GLAD nomenclature mapping files [13]. This effort by the Federal LCA Commons to develop the FEDEFL toolbox has also improved the ability of management of nomenclatures and mapping files and increased access to such resources through the publicly available github repositories and version control of lists.

Author Contributions: Conceptualization: W.W.I. and A.N.E. Software: B.Y., W.W.I. and A.N.E. Data Curation: A.N.E., B.Y., S.C. and W.W.I. Investigation: A.N.E., B.Y., S.C. and W.W.I. Methodology: A.N.E., S.C., B.Y. and W.W.I. Supervision: W.W.I. and S.C. Validation: A.N.E., B.Y. and W.W.I. Visualization: A.N.E. and S.C. Writing: A.N.E., W.W.I., B.Y. and S.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the USEPA's Sustainable and Healthy Communities Research Program. This research was supported through USEPA contract EP-C-16-015, Task Order 68HERC19F0292 with Eastern Research Group (ERG).

Institutional Review Board Statement: The U.S. Environmental Protection Agency, through its Office of Research and Development, funded and conducted the research described herein under an approved Quality Assurance Project Plan (K-LRTD-0030017-QP-1-3). It has been subjected to the Agency's peer and administrative review and has been approved for publication as an EPA document. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

Data Availability Statement: Supporting analyses can be found at <http://doi.org/10.23719/1527971>.

Conflicts of Interest: The authors have no conflict of interest.

References

1. Edelen, A.; Hottle, T.; Cashman, S.; Ingwersen, W. *The Federal LCA Commons Elementary Flow List: Background, Approach, Description and Recommendations for Use*; U.S. Environmental Protection Agency: Washington, DC, USA, 2019. Available online: https://cfpub.epa.gov/si/si_public_record_Report.cfm?dirEntryId=347251&Lab=NRMRL (accessed on 14 September 2021).
2. Suh, S.; Leighton, M.; Tomar, S.; Chen, C. Interoperability Between Ecoinvent Ver. 3 and US LCI Database: A Case Study. *Int. J. Life Cycle Assess.* **2016**, *21*, 1290–1298. [CrossRef]
3. Wolf, M.-A. *National LCA Databases: Status and Ways towards Interoperability*; Maki Consulting: Berlin, Germany, 2014. Available online: https://maki-consulting.com/wp-content/uploads/2013/05/2014_National-LCA-databases_White-paper_maki-Consulting_final_clean.pdf (accessed on 25 July 2022).
4. Janowicz, K.; Krisnadhi, A.A.; Hu, Y.; Suh, S.; Weidema, P.; Rivala, B.; Tiv, J.; Meyer, D.E.; Hitzler, P.; Ingwersen, W.; et al. A Minimal Ontology Pattern for Life Cycle Assessment Data. Available online: <https://citeseerx.ist.psu.edu/viewdoc/citations?doi=10.1.1.720.4939> (accessed on 31 August 2021).
5. Ingwersen, W.W.; Hawkins, T.R.; Transue, T.R.; Meyer, D.E.; Moore, G.; Kahn, E.; Arbuckle, P.; Paulsen, H.; Norris, G.A. A New Data Architecture for Advancing Life Cycle Assessment. *Int. J. Life Cycle Assess.* **2015**, *20*, 520–526. [CrossRef]

6. Goedkoop, M. GLAD Is Great: About the Global LCA Access to Data Initiative. Available online: <https://pre-sustainability.com/articles/about-the-global-lac-access-to-data-initiative-sustainable-data-alignment/> (accessed on 31 August 2021).
7. Kuczenski, B.; Davis, C.B.; Rivela, B.; Janowicz, K. Semantic Catalogs for Life Cycle Assessment Data. *J. Clean. Prod.* **2016**, *137*, 1109–1117. [CrossRef]
8. Kalverkamp, M.; Helmers, E.; Pehlken, A. Impacts of Life Cycle Inventory Databases on Life Cycle Assessments: A Review by Means of a Drivetrain Case Study. *J. Clean. Prod.* **2020**, *269*, 121329. [CrossRef]
9. Ingwersen, W.W. Test of US Federal Life Cycle Inventory Data Interoperability. *J. Clean. Prod.* **2015**, *101*, 118–121. [CrossRef]
10. Edelen, A.; Ingwersen, W.; Rodríguez, C.; Alvarenga, R.; de Almeida, A.; Wernet, G. Critical Review of Elementary Flows in LCA Data. *Int. J. Life Cycle Assess.* **2018**, *23*, 1261–1273. [CrossRef] [PubMed]
11. Beaufort-Langeveld, A.; Bretz, R.; Hischier, R.; Huijbregts, M.; Jean, P.; Tanner, T.; van Hoof, G. (Eds.) *Code of Life-Cycle Inventory Practice*; Society of Environmental Toxicology and Chemistry: Pensacola, FL, USA, 2003; ISBN 1-880611-58-9. Available online: <https://www.setac.org/store/viewproduct.aspx?id=1032222> (accessed on 4 June 2021).
12. UNEP/SETAC Life Cycle Initiative, GLAD: Global LCA Data Access Network. Available online: <https://www.globallcadataaccess.org/> (accessed on 11 October 2019).
13. Vadenbo, C.; Sonderegger, T.; Shobatake, K.; Selim, K.; Fazio, S.; Valente, A.; Edelen, A. Deliverable 1.3 of the Nomenclature Working Group—Methodology and Common Issues Report V2.2; UNEP Global LCA Data Access. 2022. Available online: https://github.com/UNEP-Economy-Division/GLAD-ElementaryFlowResources/raw/master/Documentation/GLAD%20EF%20mapping_Method%26common%20issues%20report_v2.2.pdf (accessed on 15 August 2022).
14. Kahn, E.; Ingwersen, W.; Skone, T.; Carpenter, A.; Wang, M.; Fox-Lent, C.; Kneifel, J.; Coventry, Z.; Arbuckle, P.; Edelen, A.; et al. Developing a Federal LCA Commons. In Proceedings of the LCA XVI, Charleston, SC, USA, 26–29 September 2016.
15. USDA; US EPA; US DOE. Cooperation on Data, Research, and Information Systems for Life Cycle Assessment/Analysis. 2018. Available online: <https://www.lcacommons.gov/> (accessed on 4 June 2021).
16. Kahn, E.; Antognoli, E.; Arbuckle, P. The LCA Commons—How an Open-Source Repository for US Federal Life Cycle Assessment (LCA) Data Products Advances Inter-Agency Coordination. *Appl. Sci.* **2022**, *12*, 865. [CrossRef]
17. US EPA Toxics Release Inventory (TRI) Program. Retrieved from TRI Listed Chemicals. Available online: <https://www.epa.gov/toxics-release-inventory-tri-program/tri-listed-chemicals> (accessed on 4 June 2021).
18. National Emissions Inventory Data. Available online: <https://www.epa.gov/air-emissions-inventories/national-emissions-inventory-nei> (accessed on 4 June 2021).
19. US EPA Resource Conservation and Recovery Act (RCRA) Regulations. Available online: <https://www.epa.gov/rcra/resource-conservation-and-recovery-act-rcra-regulations#haz> (accessed on 4 June 2021).
20. US EPA National Pollutant Discharge Elimination System. Available online: <https://www.epa.gov/npdes> (accessed on 4 June 2021).
21. US EPA Emissions & Generation Resource Integrated Database. Available online: <https://www.epa.gov/energy/emissions-generation-resource-integrated-database-egrid> (accessed on 4 June 2021).
22. USDA Agricultural Marketing Service. Available online: <https://www.ams.usda.gov/datasets/pdp> (accessed on 4 June 2021).
23. USGS National Minerals Information Center. Available online: <https://www.usgs.gov/centers/nmic/mineral-commodity-summaries> (accessed on 4 June 2021).
24. USGS. USGS Water Data for the Nation. Available online: <https://waterdata.usgs.gov/nwis> (accessed on 4 June 2021).
25. Argonne National Laboratory. The Greenhouse Gases, Regulated Emissions, and Energy Use in Technologies Model. Available online: <https://greet.es.anl.gov/> (accessed on 4 June 2021).
26. NREL U.S. Life Cycle Inventory Database. Available online: <https://www.nrel.gov/lci/> (accessed on 4 June 2021).
27. Ingwersen, W.; Ghosh, T.; Hottle, T.; Carpenter, A.; Richa, K.; Hawkins, T. Automating the Creation & Updating of Unit Process Datasets: U.S. Regional Electricity LCI Data for the Federal LCA Commons. In Proceedings of the LCA XVIII, Fort Collins, CO, USA, 25–27 September 2018.
28. GreenDelta. *openLCA*; Berlin, Germany, 2020. Available online: <https://www.openlca.org/download/> (accessed on 4 June 2021).
29. Srocka, M.; Cashman, S.; Edelen, A.; Young, B.; Ciroth, A.; Ingwersen, W. Flow Mapping in openLCA. Available online: https://cfpub.epa.gov/si/si_public_record_report.cfm?Lab=CESER&dirEntryId=349051 (accessed on 27 September 2021).
30. Young, B.; Srocka, M.; Ingwersen, W.; Morelli, B.; Cashman, S.; Henderson, A. LCIA Formatter. *JOSS* **2021**, *6*, 3392. [CrossRef] [PubMed]
31. Bare, J. The Tool for the Reduction and Assessment of Chemical and Other Environmental Impacts. *Clean Technol. Environ. Policy* **2011**, *13*, 687–693. [CrossRef]
32. Huijbregts, M.; Steinmann, Z.; Elshout, P.; Stam, G.; Verones, F.; Vieira, M.; Hollander, A.; Zijp, M.; van Zelm, R. ReCiPe 2016 V1.1 A Harmonized Life Cycle Impact Assessment Method at Midpoint and Endpoint Level Report I: Characterization. *Int. J. Life Cycle Assess.* **2017**, *22*, 138–147. [CrossRef]
33. Bulle, C.; Margni, M.; Patouillard, L.; Boulay, A.; Bourgault, G.; De Bruille, V.; Jolliet, O. Impactworld+: A Globally Regionalized Life Cycle Impact Assessment Method. *Int. J. Life Cycle Assess.* **2019**, *24*, 1653–1674. [CrossRef]
34. Young, B.; Ingwersen, W.; Bergmann, M.; Hernandez-Betancur, J.D.; Ghosh, T.; Bell, E.; Cashman, S. A System for Standardizing and Combining U.S. Environmental Protection Agency Emissions and Waste Inventory Data. *Appl. Sci.* **2022**, *12*, 3447. [CrossRef] [PubMed]

35. US EPA Terminology Services. Available online: https://iaspub.epa.gov/sor_internet/registry/termreg/searchandretrieve/termsandacronyms/search.do (accessed on 4 June 2021).
36. William, A.; Grulke, C.; Edwards, J.; McEachran, A.; Mansouri, K.; Baker, N.; Shah, I.; Wambaugh, J.; Judson, R.; Richard, A. The CompTox Chemistry Dashboard: A Community Data Resource for Environmental Chemistry. *J. Cheminform.* **2017**, *9*, 61. [[CrossRef](#)]
37. Yang, Y.; Ingwersen, W.; Hawkins, T.; Srocka, M.; Meyer, D. USEEIO: A New and Transparent United States Environmentally-Extended Input-Output Model. *J. Clean. Prod.* **2017**, *158*, 308–318. [[CrossRef](#)] [[PubMed](#)]
38. IDEA. IDEA LCA Database Home. Available online: <http://www.idea-lca.jp/> (accessed on 13 September 2021).
39. Ecoinvent. Available online: <https://www.ecoinvent.org/> (accessed on 13 September 2021).