Proceedings of the International Symposium on Sustainable Systems and Technologies, v3 (2015)

Antelope – a web service for publishing life cycle models and results

Brandon Kuczenski University of California, Santa Barbara, bkuczenski@ucsb.edu
Sabina Beraha University of California, Santa Barbara

Abstract. Preparing a life cycle assessment (LCA) requires collecting vast and varied information about industrial processes, known collectively as life cycle inventory (LCI) data. Inventory data for a given study may include a mix of proprietary information, literature results, and data sets drawn from reference databases as well as direct observations, statistical models, and estimates. Exchange and cooperative review of LCI models is challenging, because the use of private data often requires that intermediate results be concealed. Variations in modeling assumptions, scope and system boundary definitions across studies make conclusions difficult to generalize. Moreover, technical differences across software systems and the lack of a data format for publishing finished models also limit the ability of LCA practitioners to disseminate their results to a broad audience. As a consequence, it is often difficult to determine the provenance of LCA results, leading to low confidence in their interpretation.

To address the need of study authors to share their results, we have developed a web service that enables the publication of LCI models and impact assessment (LCIA) results while protecting the privacy of source data. Models are constructed out of "fragments," acyclic directed graphs in which nodes correspond to unit processes and edges correspond to intermediate flows. Nodes are weighted according to their activity levels in a manner consistent with widespread LCA practice. Fragments can be arranged hierarchically by a study author to describe complex models. Using the service, processes and fragments can be inspected, and LCIA results for both processes and fragments can be retrieved for LCIA methods that are implemented by the service. Individual model components (processes, flows, flow properties, and LCIA methods) are referenced to source files formatted according to the ILCD standard. Processes originating from datasets marked as "private" report only aggregate LCIA results, while non-private processes can be inspected at the flow level.

The service permits remote users to retrieve modeling information and LCIA results from any web terminal. The work product is designed according to the REST architectural style, meaning that queries to the service are easily intelligible, stateless, and return self-documenting and human-readable results to a user. The results can also be processed by a specialized "front-end" tool and displayed interactively. The API specification and source code are available from the sponsor agency and will be released under an open source license.

Proceedings of the International Symposium on Sustainable Systems and Technologies (ISSN 2329-9169) is published annually by the Sustainable Cognoscente Network. Jun-Ki Choi and Annick Anctil, co-editors 2015. ISSSTNetwork@gmail.com.

Cite as:
Antelope – a web service for publishing life cycle models and results. *Proc. ISSST*, Brandon Kuczenski, Sabina Beraha. http://dx.doi.org/10.6084/m9.figshare.1510989 v3 (2015)

Antelope – a web service for publishing life cycle models and results

Introduction. The LCA community has a number of software tools available for designing unit process and life cycle inventory models and performing LCA computations. These tools include high-value commercial software packages such as GaBi, SimaPro, and Umberto; lower-cost specialized tools such as MiLCA, and free systems such as OpenLCA, Brightway, and CMLCA. To perform an LCA, a practitioner must carefully construct a model in one of these tools or in a scientific computing environment, and then extract results from the model, and then interpret and report them.

One thing the community lacks is a straightforward way to publish LCA results. The traditional publication format for an LCA is a technical report describing the model, often accompanied by a lengthy supplementary file containing raw inventory or LCIA results. These documents cannot be easily read by a computer. If another researcher wishes to make use of the results to compare them against another study, or investigate the model's sensitivity to input parameters, that researcher must hope that the required information is provided by the author, and then undertake a painstaking process of interpretation in order to replicate the model.

The Antelope project originated in a desire for LCA study authors to be able to publish life cycle models in a structured way and enable readers to interactively explore the models and results. Given a product system model for a completed life cycle study, the project has 4 objectives:

1. To document the structure of the model, including data sources, material flow models, system boundary selection, and other modeling details in a structured format;
2. To dynamically compute LCIA results for unit processes, inventory modeling components, and the full product system model;
3. To provide contribution and sensitivity analysis of model results at different hierarchical levels within the product model;
4. To enable scenario analysis of models by people other than the model author, so that readers can test hypotheses about a model without the author's involvement.

The project has been implemented as a web service that provides a standard interface for retrieving model details and performing computations. The web service is paired with a client program that enables users to view the results of web service queries interactively, and to perform scenario and sensitivity analysis.

Project Origins. In 2009 the California Senate passed a law, known as SB 546, which changed how the state regulated the management of used motor oil. In order to support future policy making, the bill required the state to perform a life cycle assessment of the used oil management system in place. That study was conducted by researchers at the University of California, Santa Barbara, in collaboration with a group of industry stakeholders, a panel of critical reviewers, and several consulting firms. The study was overseen by the California Department of Resources Recycling and Recovery (CalRecycle). The study resulted in a number of technical documents available online [1, 2, 3, 4].

Upon completion in July of 2013, it was desirable for the resulting technical report to be accompanied by an interactive version of the inventory model that was created for the study. However, existing technologies fell short of the agency's requirements for transparency, cost, and ease of use. A new project was conceived to enable the study's results to be published openly and at zero cost to data users and members of the public. That project (DRR-13026) began in September, 2013 and concludes in September, 2015. "Antelope" is the nickname of the deliverable software elements and specifications for that project.

Project Components. The project was implemented as a set of software systems for obtaining information about the life cycle inventory model to be published, providing model information and impact assessment results to a remote user, and providing the remote user with an interface for interacting with the data.

The implementation has three components:

- A Data Loader that interprets LCA data provided by a study author. The data loader has two components, the ILCD Importer for main LCI data and the Fragment Importer for product system models. The data are stored in a relational database that can be accessed by the web service described below.

- A REST-style Web Service [5] (also known as the "back end") that implements the API. This component provides access to the background data, model structure, LCIA methods, and numerical results of the model in a structured format. The web service also allows users to parameterize the model, unit processes, and LCIA methods and retrieve analytical results.

- A browser-based Client Application (or "front end") to visualize models and results and allow users to interact with them in an intuitive way. The current application was designed around the anticipated needs of the original Used Oil LCA project. However, other clients can easily be developed to use the information provided via the web service.

Figure 1 shows the data flow for the tool.

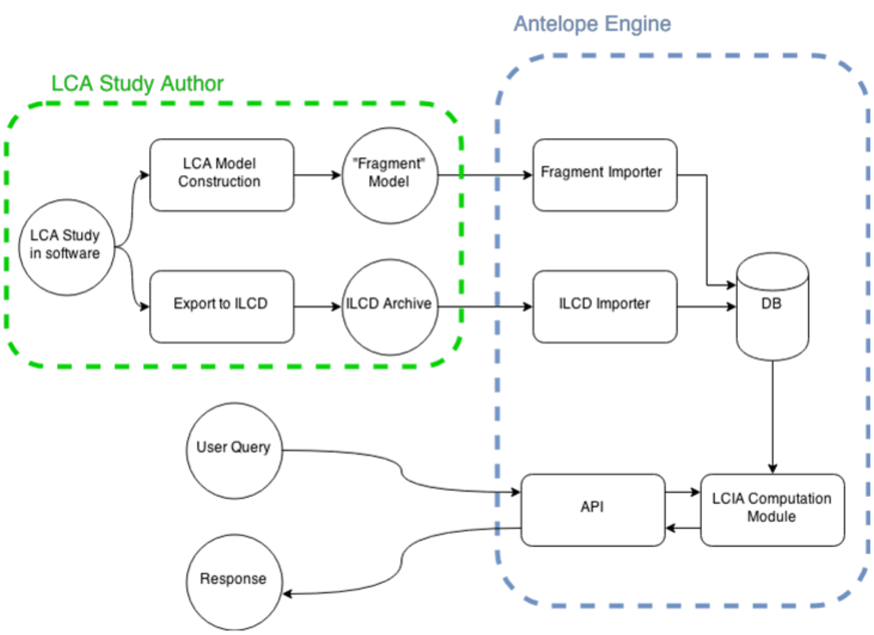

Figure 1. Back-end Data Flow. The data loader and web service are included in the blue box labeled "Antelope Engine." The client application, which interprets and presents data provided over the web API, is not shown. The data model described by the API is visualized in Figure 3.

Computational Model. To meet the project objectives it was necessary to develop a format for formally describing the product system model used to generate a life cycle inventory. Although a few serialization formats exist for life cycle inventory data [6, 7], none of them include a description of product systems. The model used for the Used Oil LCA study was the starting point for project development. It was observed that the model could be described as a directed acyclic graph and that the graph could be further broken down into a collection of nested trees arranged hierarchically. This structure was selected as the modeling framework [8].

A life cycle inventory "fragment" refers to a subgraph of an inventory model that can be represented as a tree. The nodes in the tree are processes, and the edges are "fragment flows," which express exchange relationships between the processes. At its simplest, a fragment flow is an observation that a flow occurred, and that it started or ended at a particular endpoint. Each fragment flow connects two nodes, called the parent node and the child node. The physical direction of the flow may be from parent to child or from child to parent, but the child node's activity is always inferred from the parent node's activity.

A fragment is constructed by designating a reference flow and a particular activity as a child node. That activity is then included within the fragment's system boundary. When a fragment flow terminates in an activity, that activity's child flows are added to the fragment, forming a tree, and terminated recursively. A child flow that terminates in an exchange ends the recursion. Construction of a fragment is complete when all child flows are terminated. The edge weights commonly associated with LCI models (also known as exchange quantities) are not part of the fragment definition. Figure 2 shows a depiction of an "anonymous" fragment, one in which the activities performed at each node are not specified explicitly.

Computation of the LCI and LCIA results for a fragment is performed via tree traversal, in which the fragment flows are followed recursively, beginning with the reference flow. During traversal, each node is visited and its activity level, or node weight, is determined based on the magnitude of the flow which it terminates. This requires first resolving each node to a specific unit process inventory, and then connecting each fragment flow to a particular exchange listed in that inventory. Individual nodes in the fragment represent unit processes, and the LCIA scores for these processes can be computed on a unit activity basis. Then the LCIA scores of the fragment can be determined quickly as the inner product of node weights with unit impact scores. For node weights w and unit impact scores s_t for LCIA method t, the fragment's unit impact score is:

$$S_t = \sum w_k \times s_{t,k}$$

Because the node weights and unit process scores for each node are provided separately, the tree structure of a fragment provides an inherent contribution analysis, which can be made hierarchical through the design and re-use of nested model components called "sub-fragments."

Input Data. The database is populated with data obtained from the study author via the data loader. The data loaded are of two forms: *core inventory data* and *fragment design data*.

Core Inventory Data. Antelope uses the ILCD format for describing core LCA data elements. The specification for this format is available in the ILCD Handbook [6, 9]. Antelope uses four of the seven ILCD resource types to encode the LCI data for models: *flowproperty*, *flow*, *lciamethod*, and *process*. All environmental impacts in a model thus originate in an elementary flow that is listed as an exchange in an ILCD process.

ILCD data archives loaded by the data loader can be marked as "private." Data for process resource types originating in a private archive are used for computation but are not made available to end-users. When querying process data, the API will hide exchange values for these data sets. When querying LCIA results, the API will hide the contributions from individual elementary flows, only reporting cumulative results. Finally, when providing access to ILCD source files, all exchanges except for the reference flow are removed prior to delivery.

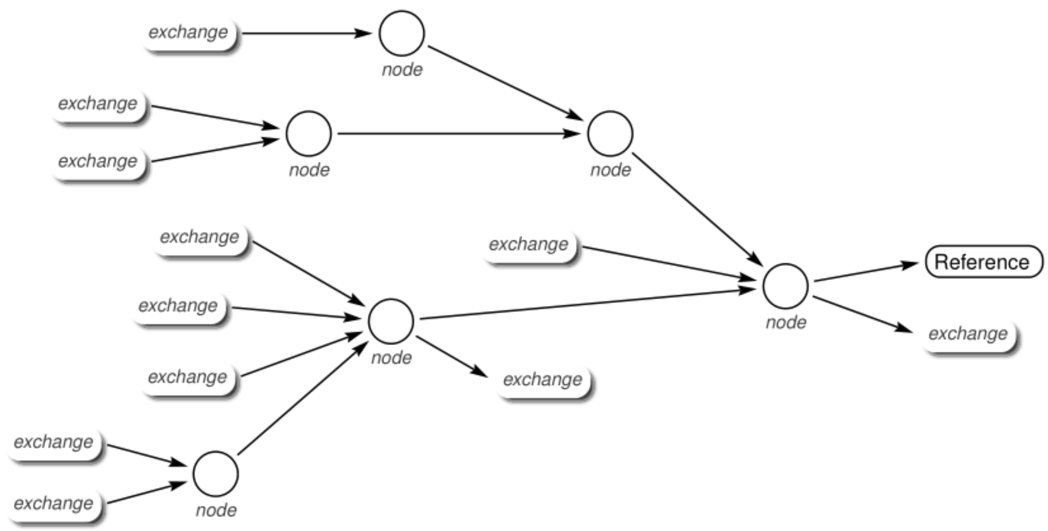

Figure 2: An anonymous fragment structure. The circular nodes represent activities occurring within the fragment's system boundary. The oval exchanges (one of which is the reference flow) represent intermediate flows that cross the fragment's boundary.

Fragment Design Data. The computation system depends on the fragment flow data model described above, which provides a standardized representation of a product system as a directed acyclic graph made up of a hierarchical arrangement of trees. The entire model structure can be represented as a single Fragment Flow table whose fields are described in Table 1 below. Each fragment has exactly one record with no parent node; this is the fragment's reference flow. The descendants of this flow (i.e. flows that are reached by traversing the fragment) make up the rest of the fragment.

Table 1. Fragment Flow data fields

Field Name	Description
FragmentFlowUUID	Unique Key for each fragment flow
ParentFragmentFlowUUID	The FragmentFlowUUID of the flow's parent node (blank for the fragment's reference flow.
FlowUUID	"Outbound" flow from the parent node
Direction	Physical direction of flow, from the parent perspective
Name	Descriptive name for the flow
Stage	life cycle stage to be used for aggregation of LCIA impacts
NodeType	Either *activity* (resolves to process or sub-fragment) or *exchange* (crosses the fragment system boundary)
TargetUUID	For *activity* nodes only, the identifier of the activity occurring at the child node, which could be a unit process or sub-fragment.

When the fragment is traversed, the exact processes or fragments used at each node are determined, along with their node weights and the magnitudes of each flow. Flows terminated as exchanges may be resolved to *background* processes. In this case, the exchange is no longer regarded to cross the fragment boundary and is instead assigned impacts within the fragment's boundary. Background resolution requires a separate table provided by the study author that relates a given flow and direction to a particular background process, whose impacts have been pre-computed and independently published.

Queries and Output Data. The API is designed to accept queries from a remote user and provide answers to those queries in a structured format. The data model is structured according to resource definitions derived from the ILCD format, but that are by extension applicable to LCA in general. The primary data resources are derived from the ILCD data types of *flow*, *flowproperty*, *lciamethod*, and *process*, to which is added *fragment*. Primary resource types have strictly metadata (i.e. documentary) information; all quantitative data in the model are associated with links between two primary data resources.

Figure 3 shows the data resources reported by the API. Quantitative data resources are shown in boxes. Several data resources are derived directly from the ILCD sources: *ProcessFlows* (pairing a process and a flow) are exchanges; *FlowPropertyMagnitudes* (pairing flow and flow property) are intermediate flow characteristics; and *LCIAFactors* (pairing a flow and an LCIA method) are LCIA characterization values. These values can be used to compute LCIA results for processes.

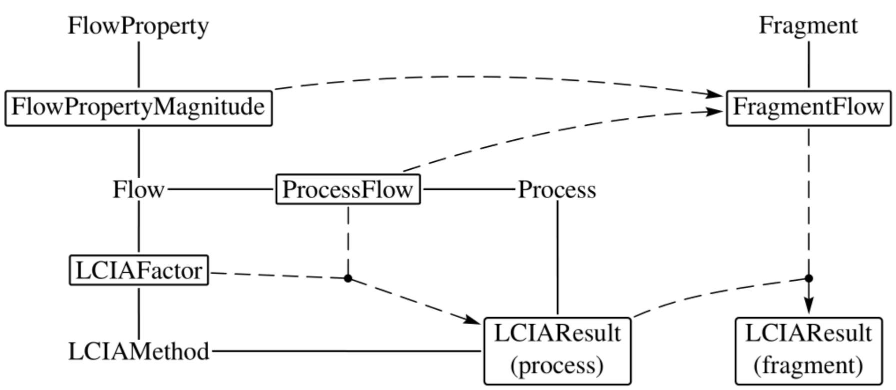

Figure 3: The API Resource Schema. Unboxed nodes are primary resource types; boxed nodes are quantitative resource types, i.e. measurements that link two primary resources; dashed arrows show the direction of information flow during computation.

ProcessFlows and *FlowPropertyMagnitudes* are used during fragment traversal to compute *FragmentFlows*, which store the node weights and flow magnitudes of fragment flows. The node weights are then paired with unit process LCIA results to compute fragment LCIA results.

Data can be queried over the Web by specifying the information desired relative to a publication URL. For instance, if the web service is published at the address http://my.example.net/antelope/, then the address http://my.example.net/antelope/api/processes/ would return a list of processes; http://my.example.net/antelope/api/processes/47 gives details about the process with ID 47; http://my.example.net/antelope/api/processes/47/processflows returns unit process exchanges; http://my.example.net/antelope/api/processes/47/lciaresults returns LCIA results for process 47; and so on.

A live implementation of the API which implements the used oil LCA model is currently running at http://publictest.calrecycle.ca.gov/lcatoolapi/api with on-line documentation available at http://publictest.calrecycle.ca.gov/lcatoolapi/help

Discussion. The web service offers a mechanism for communication of LCI models and LCIA results that is improved over current methods because it includes a structured inventory model format and supports interactive support for model interpretation and analysis. Table 2 summarizes the differences in data user experience for the different publishing mechanisms of an ISO-style report, a data archive (such as ILCD or Ecospold), and a model published using an Antelope-style web service.

Table 2. Comparison Between LCA Publishing Methods

Metric	ISO Report	Data Archive	Web Service
LCI Model Description	Interpretive (drawings and text)	Not supported	Structured (fragment)
LCIA results	Static, as-reported	Dynamic (performed by user)	Dynamic (performed by web service)
Contribution Analysis	Static, as-reported	Data user must replicate LCI model	Static, hierarchical
Scenario Analysis	Static, as-reported	Data user must replicate LCI model	Interactive, open-ended
Sensitivity Analysis	Static, as-reported	Data user must replicate LCI model	Interactive, open-ended
Result Validation	Performed by user	Performed by user	Can be performed by a validation service

The tool's capabilities support interpretation of published LCAs by permitting data users to perform independent review and test hypotheses about the results through interactive analysis. For instance, given a study showing the toxicological impacts of two different fuel combustion systems, a data user may hypothesize that differences between the systems may result from the selection of inappropriate characterization factors for heavy metals. In a static report, the user would have to rely on the study author reporting the information necessary to evaluate the hypothesis; however, using an interactively published model, the data user may directly inspect the emission factors and characterization factors used.

Web services also lend themselves well to automated data validation because source data are provided in a structured and consistent format using a standard Internet communication protocol. This facilitates the development of validation services that obtain data directly from the study to be validated and methodically compare it to data obtained from a trusted source.

Conclusion. Antelope is intended to provide an open cyberinfrastructure to enable widely distributed publication of LCI data and LCA studies. The software is based on the fragment data model for describing product systems in LCA. This data model fills a crucial gap in current LCI publishing formats in a manner that is simple, extensible, and easily adapted to existing software and data systems. The API provides a provenance framework for LCA results and enables documentation, external validation, and interpretive support of LCA studies.

Acknowledgments. Development of the fragment model and the Antelope API has been supported by CalRecycle (Contract #DRR-13026). The CalRecycle client tool was created by Sabina Beraha. Rachel Scanlon and Sabina Beraha contributed to backend development. Many thanks to Roland Geyer, Krzysztof Janowicz, Pascal Lesage, Robert Carlson, and Kyle Meisterling for support and intellectual development.

References.

[1] Roland Geyer, Ashley Henderson, Brandon Kuczenski, and Trevor Zink. Life Cycle Assessment of Used Oil Management in California pursuant to Senate Bill 546 (Lowenthal). http://www.calrecycle.ca.gov/publications/Detail.aspx?PublicationID=1465, 2013.

[2] Kline & Company, Inc. Lubricant Consumption and Used Oil Generation in California: A Segmented Market Analysis. http://www.calrecycle.ca.gov/publications/Detail.aspx?PublicationID=1511, 2013.

[3] ICF Consulting Services, L.L.C. Final Report on Direct Impacts Model (DIM) Analysis of the California Used Oil Market. http://www.calrecycle.ca.gov/publications/Detail.aspx?PublicationID=1467, 2013.

[4] ICF Consulting Services, L.L.C. Cost-Benefit Analysis and Distributional Impacts of Used Oil Management Policy Scenarios. http://www.calrecycle.ca.gov/publications/Detail.aspx?PublicationID=1466, 2013.

[5] Fielding, R. T. and Taylor, R. N. Principled design of the modern web architecture. ACM Trans. Internet Technol., 2(2):115–150. 2002.

[6] Institute for Environment and Sustainability. International reference life cycle data system (ILCD) – documentation of LCA data sets. Technical report, European Commission - Joint Research Centre, 2011.

[7] Meinshausen, I., Müller-Beilschmidt, P., and Viere, T. The EcoSpold 2 format—why a new format? Int J. Life Cycle Assess. doi: 10.1007/s11367-014-0789-z. 2014.

[8] Brandon Kuczenski. Life cycle fragments: computational methods for analyzing and sharing life cycle assessment models. In Daniel P. Ames, Nigel W.T. Quinn, and Andrea E. Rizzoli, eds. Proceedings of the 7th Intl. Congress on Environmental Modelling and Software, San Diego, CA. 2014.

[9] Marc-Andree Wolf, Clemens Düpmeier, and Oliver Kusche. The international reference life cycle data system (ILCD) format–basic concepts and implementation of life cycle impact assessment (LCIA) method data sets. In Proceedings of the 25th EnviroInfo. Conference – Innovations in Sharing Environmental Observations and Information. Schaker-Verlag, 2011.