

# **The Open Sustainability Project: a Linked Data Approach to LCA**

**Presented by Bianca Sayan, Candidate for MES, University of Waterloo**



**Massachusetts  
Institute of  
Technology**

**Leonardo Bonanni  
Hiroshi Ishii  
Chris Csikszentmihalyi**



**Hannes Ebner  
Jorge Zapico  
Marko Turpeinen  
Nils Brandt**



**Bianca Sayan  
Steven Young**



**Matthew Hockenberry**

The background is a whiteboard covered in light gray hand-drawn sketches and mathematical formulas. The sketches include a lightbulb, a person's head, a plant, a mountain, a water drop, a car, a factory, a cloud, a book, a tree, and a person sitting at a desk. The formulas include  $2+28$ ,  $640 \times$ ,  $4 \text{ years}$ ,  $71400 \text{ m}$ ,  $3100/\text{kg}$ ,  $\frac{4}{28} + \frac{1x}{140}$ ,  $X?$ ,  $X \cdot y \rightarrow ?$ ,  $\text{CO}_2$ ,  $\text{H}_2\text{O}$ ,  $\text{O}_2$ ,  $\text{me}$ ,  $\text{kg}$ ,  $\text{km}$ ,  $\text{m/s}$ ,  $\text{m/s}^2$ ,  $\text{m/s}^3$ ,  $\text{m/s}^4$ ,  $\text{m/s}^5$ ,  $\text{m/s}^6$ ,  $\text{m/s}^7$ ,  $\text{m/s}^8$ ,  $\text{m/s}^9$ ,  $\text{m/s}^{10}$ ,  $\text{m/s}^{11}$ ,  $\text{m/s}^{12}$ ,  $\text{m/s}^{13}$ ,  $\text{m/s}^{14}$ ,  $\text{m/s}^{15}$ ,  $\text{m/s}^{16}$ ,  $\text{m/s}^{17}$ ,  $\text{m/s}^{18}$ ,  $\text{m/s}^{19}$ ,  $\text{m/s}^{20}$ ,  $\text{m/s}^{21}$ ,  $\text{m/s}^{22}$ ,  $\text{m/s}^{23}$ ,  $\text{m/s}^{24}$ ,  $\text{m/s}^{25}$ ,  $\text{m/s}^{26}$ ,  $\text{m/s}^{27}$ ,  $\text{m/s}^{28}$ ,  $\text{m/s}^{29}$ ,  $\text{m/s}^{30}$ ,  $\text{m/s}^{31}$ ,  $\text{m/s}^{32}$ ,  $\text{m/s}^{33}$ ,  $\text{m/s}^{34}$ ,  $\text{m/s}^{35}$ ,  $\text{m/s}^{36}$ ,  $\text{m/s}^{37}$ ,  $\text{m/s}^{38}$ ,  $\text{m/s}^{39}$ ,  $\text{m/s}^{40}$ ,  $\text{m/s}^{41}$ ,  $\text{m/s}^{42}$ ,  $\text{m/s}^{43}$ ,  $\text{m/s}^{44}$ ,  $\text{m/s}^{45}$ ,  $\text{m/s}^{46}$ ,  $\text{m/s}^{47}$ ,  $\text{m/s}^{48}$ ,  $\text{m/s}^{49}$ ,  $\text{m/s}^{50}$ ,  $\text{m/s}^{51}$ ,  $\text{m/s}^{52}$ ,  $\text{m/s}^{53}$ ,  $\text{m/s}^{54}$ ,  $\text{m/s}^{55}$ ,  $\text{m/s}^{56}$ ,  $\text{m/s}^{57}$ ,  $\text{m/s}^{58}$ ,  $\text{m/s}^{59}$ ,  $\text{m/s}^{60}$ ,  $\text{m/s}^{61}$ ,  $\text{m/s}^{62}$ ,  $\text{m/s}^{63}$ ,  $\text{m/s}^{64}$ ,  $\text{m/s}^{65}$ ,  $\text{m/s}^{66}$ ,  $\text{m/s}^{67}$ ,  $\text{m/s}^{68}$ ,  $\text{m/s}^{69}$ ,  $\text{m/s}^{70}$ ,  $\text{m/s}^{71}$ ,  $\text{m/s}^{72}$ ,  $\text{m/s}^{73}$ ,  $\text{m/s}^{74}$ ,  $\text{m/s}^{75}$ ,  $\text{m/s}^{76}$ ,  $\text{m/s}^{77}$ ,  $\text{m/s}^{78}$ ,  $\text{m/s}^{79}$ ,  $\text{m/s}^{80}$ ,  $\text{m/s}^{81}$ ,  $\text{m/s}^{82}$ ,  $\text{m/s}^{83}$ ,  $\text{m/s}^{84}$ ,  $\text{m/s}^{85}$ ,  $\text{m/s}^{86}$ ,  $\text{m/s}^{87}$ ,  $\text{m/s}^{88}$ ,  $\text{m/s}^{89}$ ,  $\text{m/s}^{90}$ ,  $\text{m/s}^{91}$ ,  $\text{m/s}^{92}$ ,  $\text{m/s}^{93}$ ,  $\text{m/s}^{94}$ ,  $\text{m/s}^{95}$ ,  $\text{m/s}^{96}$ ,  $\text{m/s}^{97}$ ,  $\text{m/s}^{98}$ ,  $\text{m/s}^{99}$ ,  $\text{m/s}^{100}$ .

# A better way?

## What can happen when a field opens up:

## WIKIPEDIA

**English**  
The Free Encyclopedia  
3 441 000+ articles

**日本語**  
フリー百科事典  
710 000+ 記事

**Deutsch**  
Die freie Enzyklopädie  
1 135 000+ Artikel

**Español**  
La enciclopedia libre  
659 000+ artículos

**Français**  
L'encyclopédie libre  
1 017 000+ articles

**Русский**  
Свободная энциклопедия  
602 000+ статей

**Italiano**  
L'enciclopedia libera  
736 000+ voci

**Português**  
A enciclopédia livre  
617 000+ artigos

**Polski**  
Wolna encyklopedia  
736 000+ haseł

**Nederlands**  
De vrije encyclopedie  
645 000+ artikelen

search • suchen • rechercher • ricerca • szukaj • 検索 • buscar • zoeken • busca • поиск • сѡк • 搜索 • cerca • сѡк • haku • пошук • keresés • hledání • căutare • ara • 찾기 • søg • serçu • بحث • cari • tìm kiếm • претпара • suk • paieška • hladať • חיפוש • търсене • جستجو • poišči • bílga

English

**Wikipedia** – Harnessing collective intelligence to develop an encyclopedia that surpasses the Encyclopedia Britannica in several years.

**Sourcemap** - Making and Visualizing supply chains for everyone



**FoldIt** – Using the problem solving skills of non-academics to solve academic problems – In this case, finding optimal protein folding structures



# Open Data

freely available to the public without restrictions to view and use

**why?** allows for more rapid progress of knowledge, tools, and results

# Linked Data

stored in a semantic, machine-readable format

Linking data together by describing relationships between objects/concepts/everything

**why?** linked data is easier to find and connects seamlessly to other data

# Collective Intelligence

proposing a problem or task that would normally be limited to private sector or academic realms to the public

**why?** the search for a solutions benefits from more contributors and different/unusual approaches



**We want sustainability  
information to be open,  
free and easy to use**

**<http://opensustainability.info>**



# **First Steps...**

## **Get data online in a machine-readable format**

The background is a light gray sketch of a complex ontology diagram. It features various nodes and relationships represented by hand-drawn symbols and text. Nodes include a lightbulb, a person's head, a plant, a mountain, a calculator, a car, a factory, a cloud, a book, a forest, a city skyline, and a circle containing the text 'data! data!'. Relationships are shown with arrows and mathematical expressions such as  $2+28$ ,  $640 \times$ ,  $73\% \approx$ ,  $40 \text{ km/h}$ ,  $34 \text{ ppm}$ ,  $4 \text{ years}$ ,  $71400 \text{ mm}$ ,  $X?$ ,  $X \cdot Y \rightarrow ?$ ,  $\frac{1x}{140}$ ,  $84 \text{ g/l}$ ,  $\text{land CO}_2 \text{ factor}$ ,  $(2) \text{ land}$ ,  $\text{emission}$ ,  $\text{KW/CO/SO}_2$ ,  $\text{O}_2$ ,  $\text{e}_1$ ,  $\text{CO}_2$ , and  $\text{KWh}$ .

# **First Steps...**

## **Create an ontology**

# For Example - Geography

## ISO 14041

### 5.3.6 Data quality requirements

***“geographical coverage: geographical area from which data for unit processes should be collected to satisfy the goal of the study (e.g. local, regional, national, continental, global);”***

## ISO/TS 14048

***“one set of data fields for Valid geography (1.1.8) as information used to describe the geographical coverage of the data (see ISO 14041:1998, 5.3.6), expressed by:***

- unlimited number of data fields for the Area name (1.1.8.1, nomenclature 7.3);***
- one data field for the Area description (1.1.8.2);***
- unlimited number of data fields for the Sites (1.1.8.3);***
- unlimited number of data fields for the Geographical Information System (GIS) reference (1.1.8.4, nomenclature 7.3);”***



# This translates to...

```
<owl:Class rdf:ID="valid_geography">
  <rdfs:label xml:lang="en-gb">valid_geography</rdfs:label>
  <rdfs:domain rdf:resource="#process_description">
</owl:Class>
<rdf:Property rdf:ID="area_name">
  <rdfs:label xml:lang="en-gb">area_name</rdfs:label>
  <rdfs:domain rdf:resource="#valid_geography">
  <rdfs:range rdf:resource="rdfs:Literal">
  <rdf:datatype rdf:resource="xsd:string">
</rdf:Property>
<rdf:Property rdf:ID="area_description">
  <rdfs:label xml:lang="en-gb">area_description</rdfs:label>
  <rdfs:domain rdf:resource="#valid_geography">
  <rdfs:range rdf:resource="rdfs:Literal">
  <rdf:datatype rdf:resource="xsd:string">
</rdf:Property>
<rdf:Property rdf:ID="sites">
  <rdfs:label xml:lang="en-gb">sites</rdfs:label>
  <rdfs:domain rdf:resource="#valid_geography">
  <rdfs:range rdf:resource="rdfs:Literal">
  <rdf:datatype rdf:resource="xsd:string">
</rdf:Property>
<rdf:Property rdf:ID="gis_reference">
  <rdfs:label xml:lang="en-gb">gis_reference</rdfs:label>
  <rdfs:domain rdf:resource="#valid_geography">
</rdf:Property>
```

# First Steps – Build a Database

**OS.I** OPEN  
SUSTAINABILITY  
(dot) INFO

bianca.sayan@gmail.com - [Logout](#)

## Process Description

Name

## Class

This should use descriptive classes like the economic activity involved, etc.

Class Name

Reference To Nomenclature

## Add Entry

## Quantitative Reference

Quantitative Reference Type

Example: Functional Unit

Quantitative Reference Name

Quantitative Reference Unit

Should be a URI that refers to a unit of measurement

Quantitative Reference Amount

Should be a numerical value

## Technology

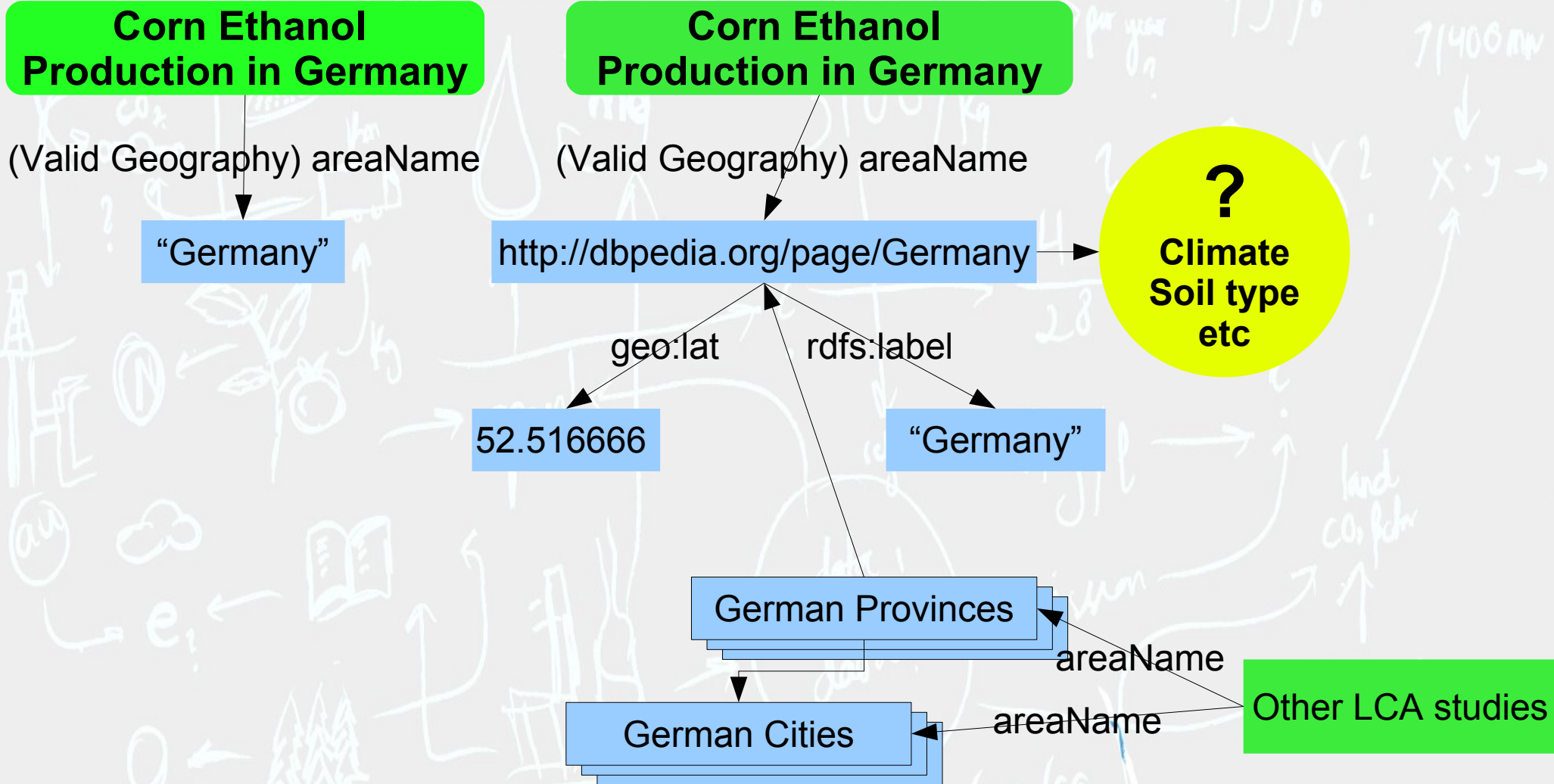
# The data will look something like this:

<code>_:YourValidGeography</code>	<code>ISO14048:areaName</code>	<code>'Germany'</code>
<code>_:YourValidGeography</code>	<code>ISO14048:areaDescription</code>	<code>"Various sites around Germany"</code>

**Which is semantic, meaning others  
can find it easily and use it!**

**But it would be a lot better if it was  
linked, too...**

# When we link to other data instead...





# First Steps – Search

Show  entries

Search:

Name	Object of Measurement	Greenhouse Gas Emissions	Energy Use	Water use
Milk	8 fluidounce of Milk	247.6 g		
Spearmint	4 g of Spearmint	259.3 g		
Cucumbers	85 g of Cucumbers	45 g		
Palm oil	1 tablespoon of Palm oil	17.6 g		
Watermelons	280 g of Watermelons	66.6 g		
Lima beans (georgia)	35 g of Lima beans (georgia)	160.7 g		
Macadamia nuts	30 g of Macadamia nuts	221.4 g		
Lamb	114 g of Lamb	1407.4 g		
Peppermint	4 g of Peppermint	285.3 g		
Plums	140 g of Plums	72.7 g		
Almonds	30 g of Almonds	166.1 g		
Tomatoes	85 g of Tomatoes	87.8 g		
Blackstrap molasses (new orleans)	1 tablespoon of Blackstrap molasses (new orleans)	2 g		
primary aluminum ingot	primary aluminum ingotkg of 1000			
Surface mounting of printed board assembly (PBA)	1 pieces of Printed board assembly (PBA)			
primary aluminum ingot	1000 kg of primary aluminum ingot		140000 kg	10900 Liters

Showing 1 to 16 of 16 entries (filtered from 99 total entries)

# Why OSI?

- **LCA-specific issues: transparency, iterative improvement, feedback/review**
- **Tie together existing databases**
- **Anyone publishes an LCA → anyone can instantly access it**
  - **Same thing for new tools, models, ontologies, and visualizations**
- **Different data formats handled seamlessly and work together**
- **Easily build on existing body of work by using and modifying open source applications and open data. Reduced redundancy.**
- **Greater interaction and mutual benefits with LCA and non-LCA community (feedback, validation, e tc)**

# Challenges

- Building a sustained community of LCA experts
- Good open and linked supplementary data sets are still rare
- Must catch up to the available functionality of existing LCA software (capability)

# Next Steps

- Building a community
- Adding data
- Make it publicly accessible
- Add other Ontologies : EcoSpold, etc
- Find or create open and linked supplemental data: NCAIS, chemical databases
- An “AppGarden”, showcasing published and available applications

More functionality: normalize and graph impacts, facilitate peer comments and validation

- Help participants submit data quickly by uploading it in existing formats



Title: Biofuels

Reference flow: kg of Ethanol

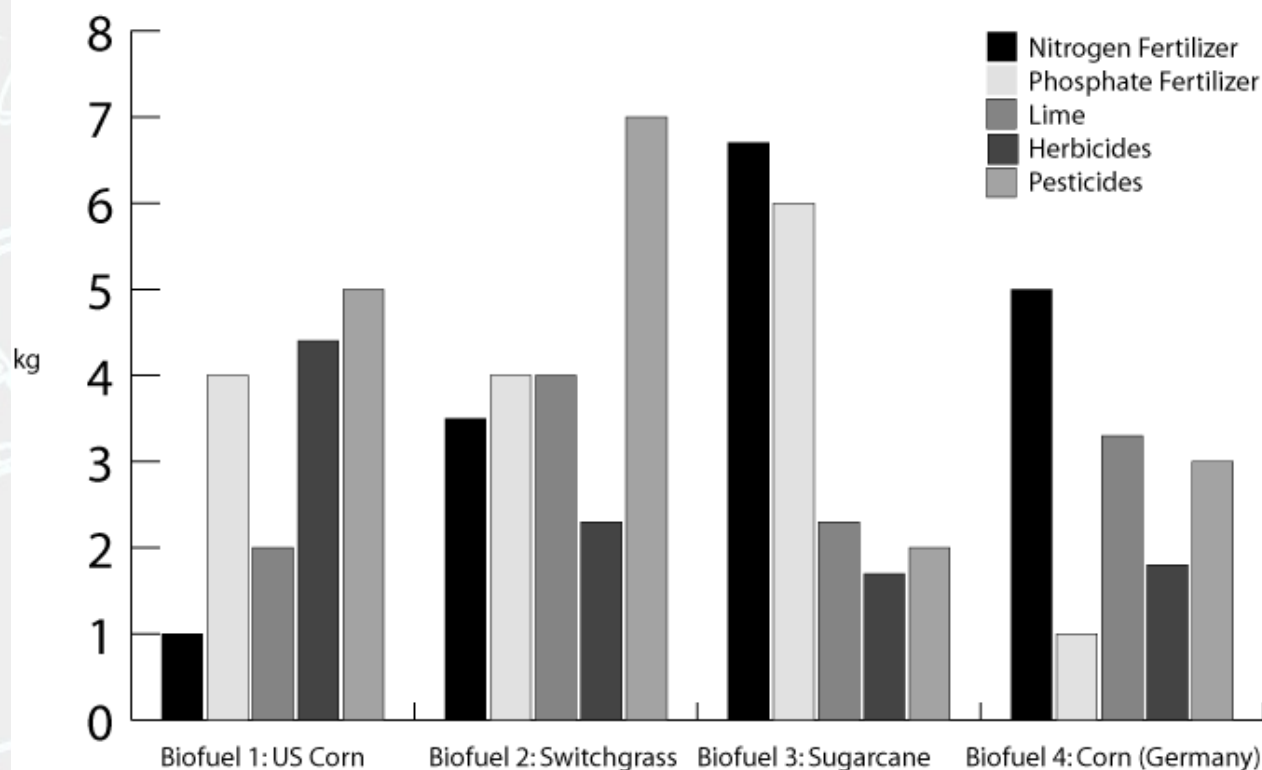
Graph Type: Inputs

### Data

<http://db.opensustainability.info/4055584322.rdf>  
<http://db.opensustainability.info/4456578129.rdf>  
<http://db.opensustainability.info/1018502335.rdf>  
<http://db.opensustainability.info/1257994368.rdf>

Add another dataset

- ☒ Nitrogen Fertilizer (same as)
- ☒ Phosphate Fertilizer (same as)
- ☒ Lime (same as)
- ☐ Potassium Fertilizer (same as)
- ☐ Potash (same as)
- ☒ Herbicides (same as)
- ☒ Pesticides (same as)
- ☐ NPK Fertilizer (same as)



### References

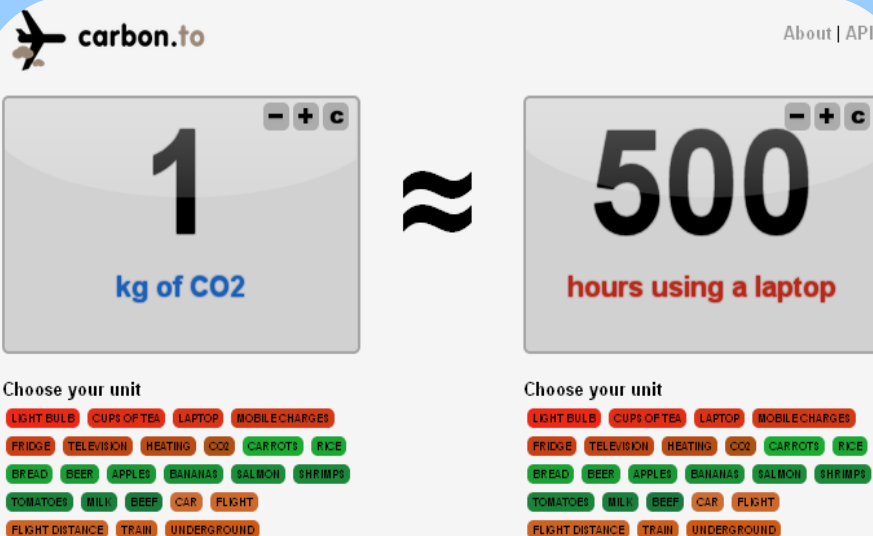
Fukushima, Y, and S P Chen. 2009. "A decision support tool for modifications in crop cultivation method based on life cycle assessment: a case study on greenhouse gas emission reduction in Taiwanese sugarcane cultivation." International Journal of Life Cycle Assessment 14:639-655.

Kim, S, and B Dale. 2005. "Environmental aspects of ethanol derived from no-tilled corn grain: nonrenewable energy consumption and greenhouse gas emissions." Biomass and Bioenergy 28:475-489.  
<http://linkinghub.elsevier.com/retrieve/pii/S0961953404001916>.

(Mock up)

# Providing Valuable Data to Important Sustainability Applications

**Sourcemap** is an online application that helps businesses, researchers, and the public create their own supply chain maps to help them measure their own environmental impact.



**Carbon.to** helps people to understand the magnitude of GHG emissions by relating them to everyday activities and objects.

The most difficult aspect of open projects are not technical in nature. An open community is nothing without participants and their contributions. Here is how to contribute:

### **Share your Data**

Go to **<http://opensustainability.info>** and start sharing

### **Or tell us why you will/won't**

Help us understand your views on data sharing by participating in our survey starting November 15:

**<http://opensustainability.info/survey/>**

This data will also be used in the SETAC workshop on LCA database standardization in early 2011



# Join Us.

**info@opensustainability.info**  
**http://opensustainability.info**



**Massachusetts  
Institute of  
Technology**

**Leonardo Bonanni  
Hiroshi Ishii  
Chris Csikszentmihalyi**



**Hannes Ebner  
Jorge Zapico  
Marko Turpeinen  
Nils Brandt**



**Bianca Sayan  
Steven Young**



**Matthew Hockenberry**