PAUL SCHERRER INSTITUT

PSI

Wir schaffen Wissen – heute für morgen

Paul Scherrer Institut

Chris Mutel

Brightway2

A new contribution to open source industrial ecology

PSI, 26. Juni 2015

Brightway2

Free and open source

In development since 2012

>12.000 sloc
>100 pages of documentation

Extensions for:
Dynamic LCA
Regionalization
Parameterization
Meta-processes

Not for everyone!

LCA

Brightway2 is open source software for IO and LCA industrial ecology calculations. It is a **framework** for calculations, and works especially well when integrated with other models or databases.
There is a lot on Brightway2 online: http://brightwaylca.org/, http://brightway2.readthedocs.org/en/latest/, so this presentation will not be extensively annotated.

Design goals: Freedom

- **No fixed** database schema

```
{
    'code': 'T17 SSAFRICA',
    'database': 'gmm-model',
    'linker': 'Background databases',
    'location': 'SSAFRICA',
    'name': 'Hydrogen Fuel Cell Transport',
    'output_name': 'T1',
    'row': 48,
    'sheet': 'demands',
    'spreadsheet name': 'Transport, fuel cell hybrid, electrolysis’,
    'technology': 'T17',
    'unit': 'meter'
}
```


One primary design goal is not to force a certain database schema; rather, the database schema should be a function of the question to be answered and the original data.
In the screenshot, the black fields are normal LCA data fields, and blue are added specifically for this example.

Design goals: Freedom

- **No fixed** database schema
- **Projects** are self-contained workspaces

```
In [1]: from brightway2 import *

In [2]: projects.current
Out[2]: 'default'

In [3]: list(databases)
Out[3]: ['biosphere3']

In [4]: projects.current = 'GMM'

In [5]: list(databases)
Out[5]:
['CARMA 2025',
 'ecoinvent 3.1 cutoff without energy',
 'ecoinvent 2.2',
 'ecoinvent 2.0',
 'ENTSO',
 'ecoinvent 3.1 cutoff',
 'biosphere3',
```


Each **project** is a subdirectory with its own copy of all data, so you can keep a copy of each project or paper, and modify copies without worrying about corrupting your primary data.

Design goals: Freedom

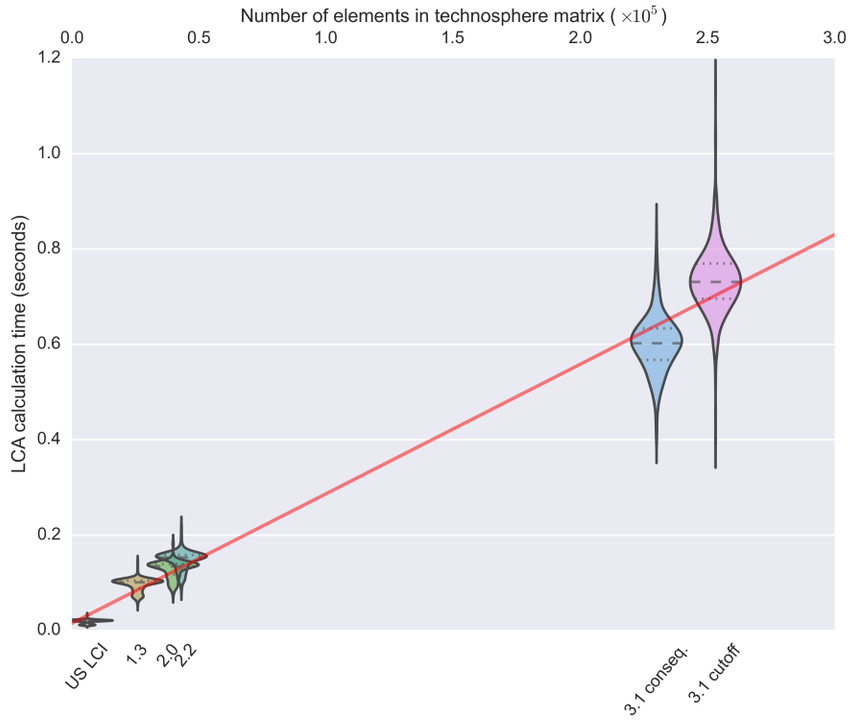
- **No fixed** database schema
- **Projects** are self-contained workspaces
- Easy **data manipulation**

```
In [1]: from brightway2 import *

In [2]: for activity in Database("ecoinvent 3.1 cutoff"):
   ...:     for exchange in activity.technosphere():
   ...:         if exchange.input['unit'] == 'kilowatt hour':
   ...:             exchange['amount'] = 0
   ...:             exchange.save()
   ...:
```


Data transforms can be written in Python, and make it easy to do data manipulation.

Design goal: Power



A final design goal is performance. Good performance of static and stochastic calculation is a prerequisite for exploring new ideas, as well as uncertainty and sensitivity analysis.

Design goal: Power

Compute LCA scores for all of ecoinvent 3.1 cutoff in < 5 minutes

- One LCIA method
- Single threaded
- 2014 Desktop

> 100 Monte Carlo iterations a second

- Multithreaded
- 2011 Labtop
- Ecoinvent 2.2

Program structure

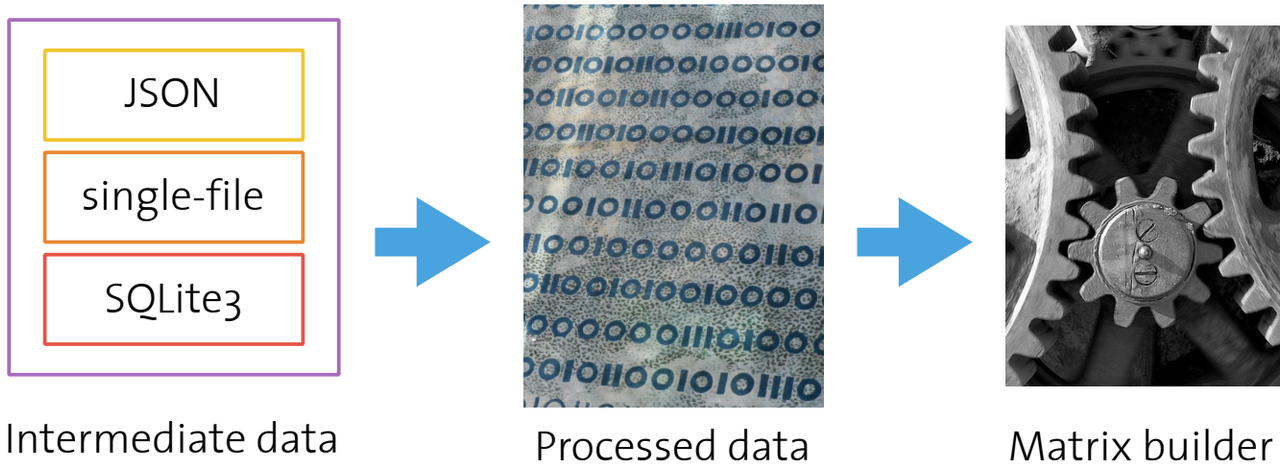


Photo credits:
https://www.flickr.com/photos/mikecogh/8192314996
https://www.flickr.com/photos/brentinoz/4221291984

Data is stored in two ways: As the raw data, with all text fields, and as a structured array with only numerical values which is faster to load and prepare for calculations.

Design goal: Integration

Python is the *de facto* language of scientific programming

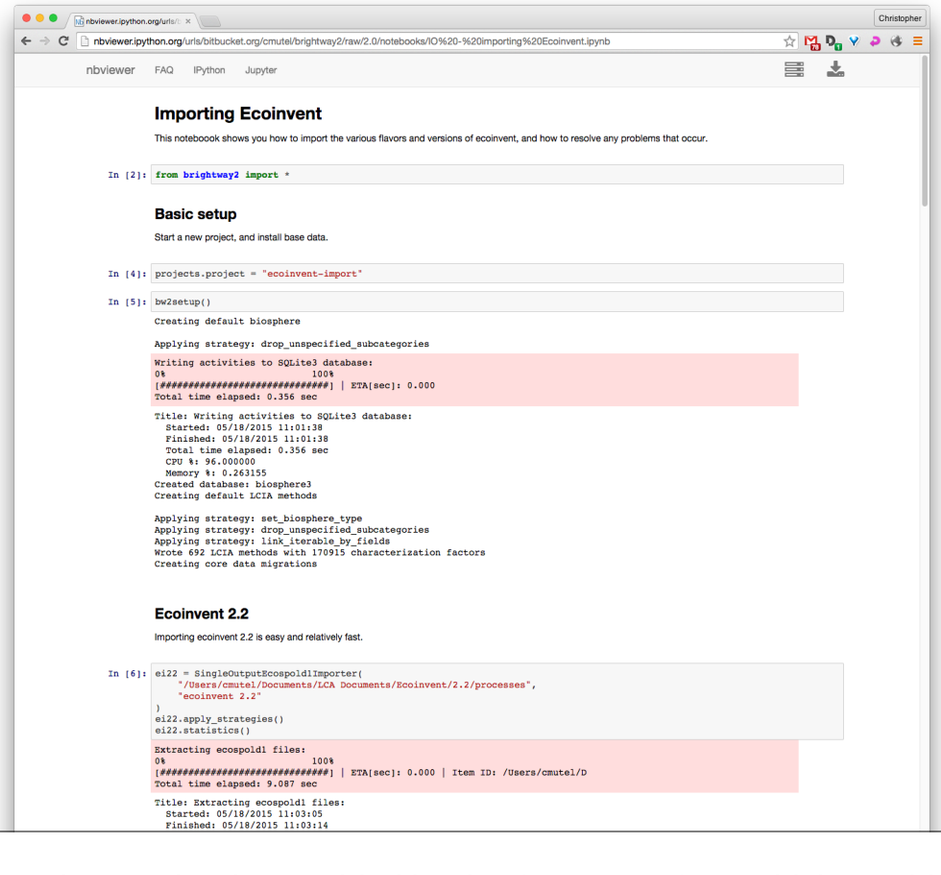
Develop and integrate your own Python model

Incredible library of python packages

Integrate with other languages

AB is the activity browser.

User interfaces



One fantastic interface for scientific work are the interactive notebooks provided by the Jupyter project, which combine working code with notes, graphics, interactive Javascript, and other awesomeness.

Version 2.0

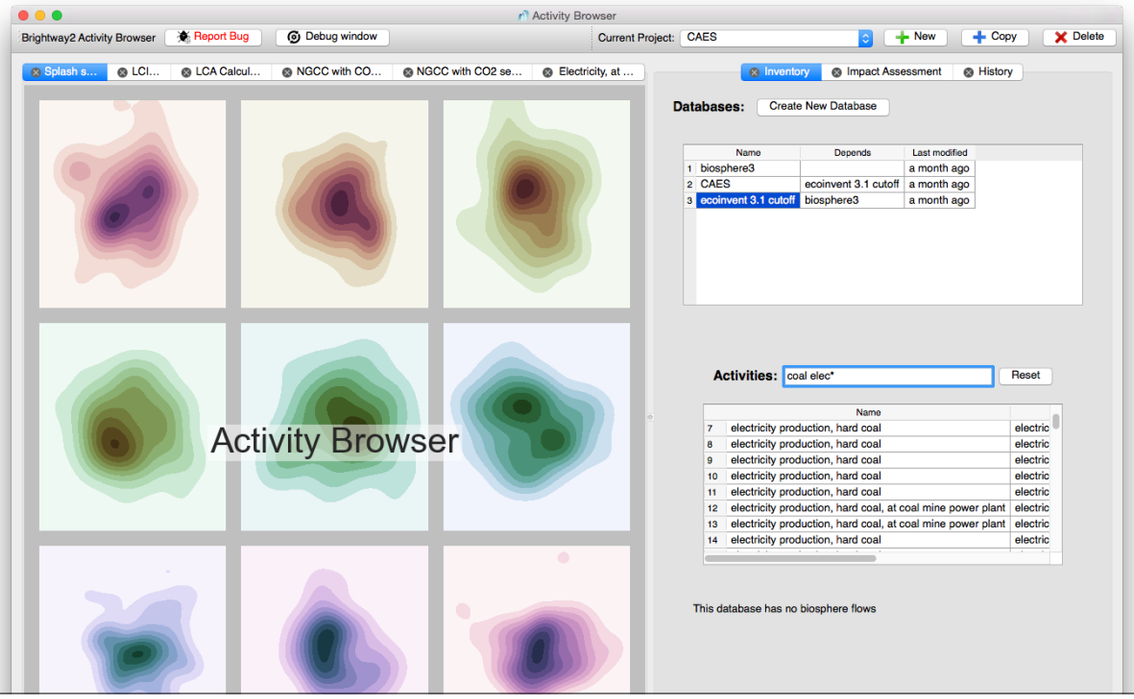
- Full **text search**

```
In [1]: from brightway2 import *

In [2]: Database("biosphere3").search("carbon dioxide foss*")
Out[2]:
['Carbon dioxide, fossil' (kilogram, None, ('air', 'non-urban air or from high stacks')),
 'Carbon dioxide, fossil' (kilogram, None, ('air', 'urban air close to ground')),
 'Carbon dioxide, fossil' (kilogram, None, ('air', 'lower stratosphere + upper troposphere')),
 'Carbon dioxide, fossil' (kilogram, None, ('air', 'low population density, long-term')),
 'Carbon dioxide, fossil' (kilogram, None, ('air',)),
 'Carbon dioxide, non-fossil' (kilogram, None, ('air', 'low population density, long-term')),
 'Carbon dioxide, non-fossil' (kilogram, None, ('air', 'non-urban air or from high stacks')),
 'Carbon dioxide, non-fossil' (kilogram, None, ('air', 'urban air close to ground')),
 'Carbon dioxide, non-fossil' (kilogram, None, ('air', 'lower stratosphere + upper troposphere
 'Carbon dioxide, non-fossil' (kilogram, None, ('air',))]
```

Version 2.0

- Full **text search**
- **Graphical** interface



Version 2.0

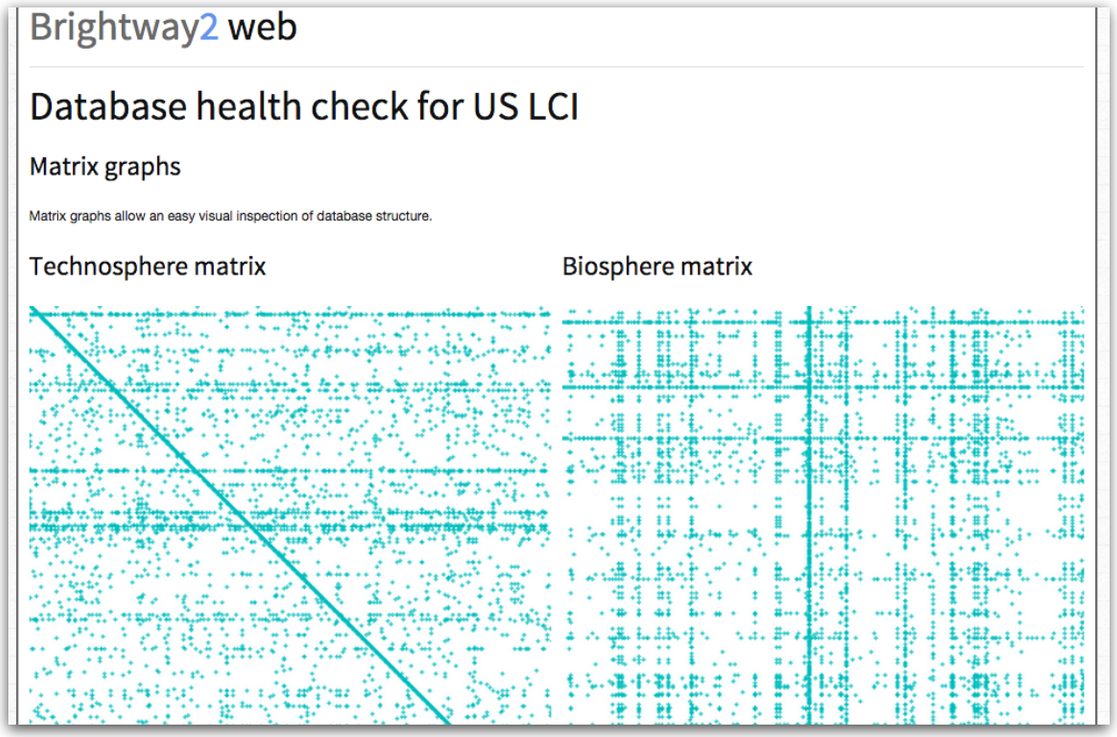
- Full **text search**
- **Graphical** interface
- Traverse supply chain in **both directions**

```
In [4]: for exc in activity.upstream():
   ...:         print(exc)
   ...:
Exchange: 0.097577 kilogram 'market for indium' (kilogram, GLO, None) to 'indium tin oxide powder production, nano
scale, for sputtering target' (kilogram, RoW, None)>
Exchange: 0.0054898 kilogram 'market for indium' (kilogram, GLO, None) to 'photovoltaic laminate production, CIS'
(square meter, RoW, None)>
Exchange: 0.022688 kilogram 'market for indium' (kilogram, GLO, None) to 'sputtering target production, sintered,
indium tin oxide' (kilogram, RER, None)>
Exchange: 0.0054898 kilogram 'market for indium' (kilogram, GLO, None) to 'photovoltaic laminate production, CIS'
(square meter, DE, None)>
Exchange: 0.022688 kilogram 'market for indium' (kilogram, GLO, None) to 'sputtering target production, sintered,
indium tin oxide' (kilogram, RoW, None)>
Exchange: 0.00089368 kilogram 'market for indium' (kilogram, GLO, None) to 'photovoltaic laminate production, a-Si
' (square meter, RoW, None)>
Exchange: 0.097577 kilogram 'market for indium' (kilogram, GLO, None) to 'indium tin oxide powder production, nano
scale, for sputtering target' (kilogram, RER, None)>
Exchange: 0.00089368 kilogram 'market for indium' (kilogram, GLO, None) to 'photovoltaic laminate production, a-Si
' (square meter, US, None)>
```


One key change in version 2.0 is the use of a database table for exchanges, which allow us to traverse the supply chain in both directions, as well as easily update exchange links.

Version 2.0

- Database **health check**



The database health check is an online report that looks at a number of different data quality indicators for a database.

Version 2.0

- Database **health check**

Multioutput processes

Multioutput processes can be used in Brightway2 (see Multioutput processes in LCA), but should be handled with care. Each multioutput process should be inspected carefully to make sure the intended effects will occur.

There are no multioutput processes in this database

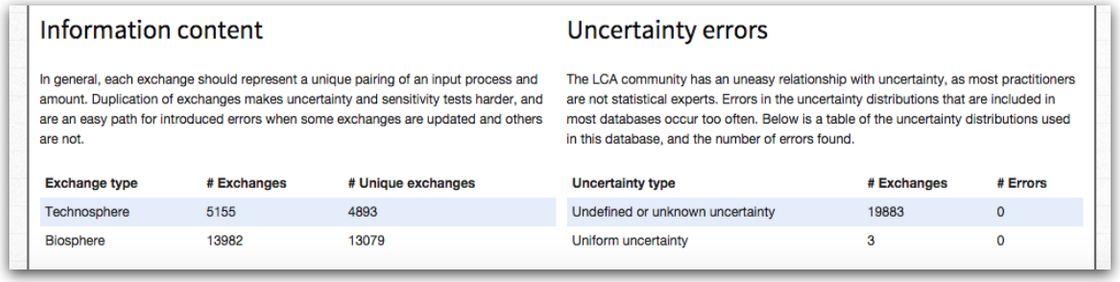
Page rank

This are the 20 most important processes, when the technosphere matrix is analyzed using the Page Rank algorithm. Page Rank scores are normalized to the number of processes in the database, i.e. each score is actually *page rank score * number of processes in database*. Anything higher 50 is pretty high - these processes should be audited for completeness and accuracy, as they will be very important in calculating impact assessment scores.

Name	Score	Unit	Categories	Location
Diesel, at refinery	98.98	L	Petroleum and Coal Products Mnf.,Petroleum Refineries	RNA
Residual fuel oil, at refinery	35.92	L	Petroleum and Coal Products Mnf.,Petroleum Refineries	RNA
Transport, combination truck, diesel powered	33.64	ton kilometer	Truck Transportation,General Freight Trucking	US
Electricity, at Grid, US, 2008	32.85	kilowatt hour	Utilities,Utilities	RNA
Transport, train, diesel powered	31.91	ton kilometer	Rail Transportation,Rail Transportation	US
Transport, combination truck, average fuel mix	30.97	ton kilometer	Truck Transportation,General Freight Trucking	US
Natural gas, combusted in industrial boiler	21.81	cubic meter	Utilities,Steam and Air-Conditioning Supply	US
Transport, barge, diesel powered	21.63	ton kilometer	Water Transportation,Inland Water Freight Transportation	US
Transport, barge, average fuel mix	21.43	ton kilometer	Water Transportation,Inland Water Freight Transportation	US
Transport, barge, residual fuel oil powered	21.27	ton kilometer	Water Transportation,Inland Water Freight Transportation	US
Transport, pipeline, unspecified petroleum products	20.05	ton kilometer	Utilities,Fossil Fuel Electric Power Generation	RNA
Liquefied petroleum gas, combusted in industrial boiler	17.37	L	Utilities,Steam and Air-Conditioning Supply	US

Version 2.0

- Database **health check**



Information content

In general, each exchange should represent a unique pairing of an input process and amount. Duplication of exchanges makes uncertainty and sensitivity tests harder, and are an easy path for introduced errors when some exchanges are updated and others are not.

Exchange type	# Exchanges	# Unique exchanges
Technosphere	5155	4893
Biosphere	13982	13079

Uncertainty errors

The LCA community has an uneasy relationship with uncertainty, as most practitioners are not statistical experts. Errors in the uncertainty distributions that are included in most databases occur too often. Below is a table of the uncertainty distributions used in this database, and the number of errors found.

Uncertainty type	# Exchanges	# Errors
Undefined or unknown uncertainty	19883	0
Uniform uncertainty	3	0

More information:

- Email: cmutel@gmail.com
- Website: brightwaylca.org
- Development blog: chris.mutel.org