

*Working together towards a common goal:*

*Open access data and open source software for transparent product footprint*

## **Agenda:**

- 1) Background and updates on recent individual developments (30 min)
- 2) Towards open access data and open source software: achieving full modularity of independent developments and parallel work, e.g. the Industrial Ecology society (45 min)
- 3) Discussing opportunity to suggest BONSAI as an add-on to UNEP's GLAD project (15 min)
- 4) Re-launching our current communication channels ([GitHub](#), [LinkedIn](#), [Working Groups](#), Board) (20 min)
- 5) AOB

### **1. Summary of recent individual activities and updates**

*Romain Sacchi:*

- Working on a simplified attempt to have an online platform that can store LCI information in one or several DBs. Admins receive requests for modification/update of LCI data from users. The platform should perform routine functions to scrap data to update LCI data (U.S GS, UN stats, Comtrade, etc.). Can also offer an interface to solve LCIs, perform LCA calculations. For now, Django and Brightway are used to do that.

*Konstantin Stadler (NTNU):*

- Finalized EXIOBASE 3
- Method paper describing EXIOBASE compilation accepted at JIE
- EXIOBASE will be open access (after embargo period), not quite clear yet about derivatives, but it will definitely be share alike.
- Continue work on pymrio - webinar about use probably in Mid January, same time a peer reviewed publication of the software is planned.

*Stefano Merciai (2.-0 LCA):*

- Working on EXIOBASE v3 Hybrid Supply-Use tables (HSUTs), where mass, energy and monetary levels are merged into one balanced framework. The algorithm is written in GAMS.
- Continuously fixing the EXIOBASE v3 HSUTs bugs.

*Tair Bilyalov (Circle Economy)*

- We are building an online dashboard that measures a city's performance in the circular economy and is comprised of a number of interactive digital gauges, some of which will employ the EXIOBASE database (material flow analysis, waste production, material requirement, ...).
- We are looking to pull input-output data for the material flow analysis (preferably on a city-level), and are at the same time interested in environmental extensions concerning waste and (embedded) emissions for other indicators.

- On the infrastructure side we have migrated EXIOBASE 3.3.11 from Excel into an online relational database and we are exploring possibilities to enable various local parties to upload, correct, or add data to keep it accurate and up to date.

*Pascal Lesage (CIRAIG)*

- Building Brightway modules that would allow the use of presampled data arrays (resulting from e.g. Monte Carlo simulations or time series) stored in external servers. This allows the creation of light tools (using e.g. aggregated data) that can still run uncertainty analyses
- Coded other Brightway modules or add-ons that allow LCA with aggregated datasets with presampled Monte Carlo samples, Global Sensitivity Analyses, adding of application uncertainty, and general tools aimed at allowing LCA of underspecified systems (quick results, guided improvement)
- Some work on openIO-Canada.

*Stefan Pauliuk (Uni Freiburg, Germany)*

- Together with Guillaume Majeau-Bettez I developed pySUT, an open source Python library for handling supply and use tables (SUTs) and converting them into IO models: <https://github.com/stefanpauliuk/pySUT>
- I curate the Indecol Dashboard on Github, an inventory of open source software available to the field: <https://github.com/IndEcol/Dashboard>
- Got funding for a PhD position for statistical modelling of hybrid supply chain models, under OASES project under leadership of Chris Mutel. Position will commence during 1st half of 2018.

*Mark Goedkoop (Pré)*

- New to Bonsai
- Used Exiobase in some tools for banking sector and IUCN for biodiversity: [www.biosope.info](http://www.biosope.info)
- Currently developing (together with Quantis and Blonk) a collaboration platform to facilitate makers of specific databases to add their data and offer it to LCA practitioners. One idea is to develop a core database with energy and transport, these specific databases can link to. Later many more data content providers shall be invited. Currently investigating how Bonsai would fit in. We would like to create the market mechanisms that make it attractive for content providers to offer their data, either on a commercial or free of charge basis.

*Tomás Navarrete Gutiérrez (Luxembourg Institute of Science and Technology)*

- My activities are more on the “user” side of bw2
- Mostly linking computational models (ABM) to LCA

*Chris Davis (University of Groningen)*

- Work on using linked data to translate between product & industry classifications ([Github repository](#), [presentation](#)). By linking together existing concordances, it's possible to then create networks of “translation pathways” that allow us to translate between two sets of classification that may not have a concordance created for them.
- Fuzzy matching on LCI process names (see example [article](#)). This relies on a machine learning technique that analyzes large amounts of text. Based on this, it can calculate a metric indicating the similarity between any two pairs of words. In the example from the

article, we demonstrate how we can use this to link a process describing “softwood” to processes describing “cedar” and “spruce”. The advantage of this approach is that these relationships are learned entirely by an algorithm, which means that you don’t have to compile a giant thesaurus by hand.

- Initial work on converting forwast (Excel files) to rdf: <https://github.com/cbdavis/forwast-to-rdf>.

Massimo Pizzol (Aalborg University)

- Developed a python class [IOready](#) to handle IO tables in exiobase format. i.e. prepare them for further calculations. I think this could be ideally combined with the other python libraries mentioned above by Konstantin and Stefan (it is not meant to substitute them...)
- I started working on possibilities for importing exiobase 3 in Brightway but got stuck due to lack of time (free time project). Would appreciate any help.
- As Tomás, “My activities are more on the “user” side of bw2”
- I would be really interested in knowing more about the “Wikipedia” concept of BONSAI. I.e. the idea of a bottom-up upload of inventories from a community of users, with peer-review of the inventories. As far as I remember this was one of the original ideas of BONSAI, but maybe I lost track of things...My questions in this respect would be: what is the status regarding this process? How do you plan to build a community? What are ideas for templates for upload of inventories? And for peer review? Any feedback with the various initiatives going on on transparency and reproducibility of LCA inventories? I think this is a very innovative idea so I hope you will pursue it. I am willing to contribute of course.
- (Sorry for missing the meeting, I am on a train...)

Your name (Your organization)

- Activities
- Updates
- ...

## 2. Towards full modularity of independent developments towards our common vision and relation to parallel work, e.g. the Industrial Ecology society

There was an overall consensus to use the BONSAI wiki as a technology roadmap. We agreed to contribute drafting this map with developed components and future ones to be prioritized and [modifying](#) the drafts of possible development stages. The content is open to changes and contribution. Everybody is welcome to contribute as preferred and/or add priorities, time and potential candidates info for each subtask (if existing, else create new one).

Engagement in contributing to the wiki is crucial and may give rise to small working groups on specific issues of interest. The *Issues* tab on GitHub can be used to address technical problems. The *Wiki* tab could be used for management purposes.

The problem of a common ontologies is central in all projects aiming at creating a semantic web of data. Chris Mutel proposed to make a list of ontologies similar to the list of software that Stefan has made on IndEcol dashboard (all links below). This may also be a topic for a smaller working group (see also table below). Tomas added that the next step to the list of available tools is to figure out

how and these tools fit. Michele pointed out that this was the main objective of having a wiki but it requires engagement and participation of the community to work.

Pascal agrees to make the wiki the place where developers can quickly find out pieces of the puzzle missing, available and how they fit together. The wiki should create human links. Bo pointed out that this what the wiki offers to do: for example, under data->data quality->uncertainty Pascal (working on uncertainty) could write his name and activities to reach out to the community. Andreas also agrees that it can well function as a repository of code and a technology root type graphic and connect to e.g. OPEN Nexus data and other available tools. The wiki should not follow a top-down approach though but a bottom up one. It could also facilitate projects funds applications.

Chris Mutel asks how could BONSAI help effectively? Why should somebody that already has a million projects also be involved in BONSAI? How do you get busy people to buy in and how can they be helped to solve their own problems? Chris opinion is that BONSAI is a tool to coordinate from the starting point decentralized projects (e.g. data formats, conventions, ontology alignments) thus making projects useful in the broader sense (modular). The wiki right now can be overwhelming because deals to many different topics.

Michele: the purpose of the BONSAI's wiki (the wiki, not the overall BONSAI project goal) was exactly to have a unique virtual discussion forum. Its large scope intended to cover all different aspects of the project under the same platform (a sort of management platform). The wiki has been moved on GitHub to create a parallel between project management with actual software development.

Pascal: the goal of the BONSAI project is more than the function of the wiki (linking ongoing projects, similarly to what the Industrial Ecology Dashboard does), because BONSAI provides an extra push to brings the pieces together to create what ends up being a db. Clearly this extra step is more difficult but it is also the stimulating part of this network.

Tomas: the added value of BONSAI is that fosters reproducibility, because it is open and makes clear from where all the pieces come or, even when the software is not clear in some parts, one can trusts the developers because of their expertise (e.g. scientific libraries and so on). BONSAI tasks is therefore about specifying the working flow allowing people to publish code next to the existing one and should not restrict the compatibility to a specific platform.

Mark: the objective of BONSAI *and deliverables* needs to be very clear, also for attracting investments. Is BONSAI a place to coordinate all kind of research? Is it a conductor of an orchestra of researchers or the composer of a new piece of music? The hesitation from the musician seems perhaps originating from the lack of clarity about the utility of joining the orchestra. We got some very useful inputs today.

Michele: it is a new piece of music (it has never done before) and is trying to offer a conductor and a stage for the musicians to sit together. This requires either funds or in kind contributions in synergies with ongoing projects. It is important to give continuity to this discussion and the work of the BONSAI network. Proposals will be advanced following this meeting.

Mark proposed to coordinate a "marketing team" for fundraising. His idea is to have a network building phase with a physical meeting to discuss a shared vision and steps to take.

Bo: the idea of the wiki (the wiki, not the whole BONSAI project) is actually to build a technology roadmap. Chris and Bo also proposed to write a scientific article describing the technology roadmap as an incentive to work on refining the objectives, technical specifications and potential deliverables of the BONSAI network.

Brandon: this is an useful conversation; he is working on inventory model (it fits in the product system model in the bonsai wiki).

Stefan: very interesting conversation. It seems clear that what is needed is leadership for coordination and prioritization of work items. The interface between different software block seems to be missing because the interfaces are not defined. Leadership may help with that too but the idea of a user interface where the software pieces are arranged together seems to be a priority. On the wish list Stefan would put a place where we can share inventory data with a few simple requirements for compliance (machine readable). Michele supports the idea and its top priority.

Andreas: This is similar but not exactly *Open Nexus* does: it compiles data sources from different providers while BONSAI aims to allow users to upload data that are then automatically integrated. His team completed the collaboration server on Open LCA which is indeed allowing to upload LCA models on a common platform (release in November). The classification for the zolca format is public and machine readable. GLAD nomenclature working group is also working on nomenclature and harmonized classification between databases too and they open to get feedbacks.

Chris D.: done some work with Brandon (see description above under point 1) trying to link together process name in LCA databases; they did not try to create a masterclass classification but leave open the possibility to create link and disambiguate. Currently we still need a human and the end answering the question: "*are this two items the same or not*". We need to speed up and automate this process. This may be a good way to get started. Technically he is using machine learning procedures that read text and is able to calculate similarities between terms. He sees it as a fuzzy-matching work since we can't anyway structure everything, thus we need procedures to translate between different nomenclatures.

Bo: it seems that there is so much great work going on. We clearly need coordination. Michele has done a lot of work on this but funds are running out and we have not been able to show yet that this work is useful to the community. How can we carry this on?

Tomas: what about organizing workshops for doing actual coding before or after conferences (SETAC or LCM for examples) and recruiting new developers? Bo supports the idea.

Mark: I hear the need for coordination and structuring individual developments but not to overplan things because we risk to kill creativity. Investors need to know for what they are paying (coordination? development?). We should start smaller and begin with the coordinating role. The challenge is to get everybody on board making it attractive for everybody to contribute.

Brandon: part of the issue may be having a bottom up perspective, figuring out what can be done that would be useful immediately, e.g. put 2 individual pieces together.

Romain: agrees with the importance of minimal, limited but concrete developments producing something that can be shown. This should come before discussion on the nomenclature and other

more complex discussion. This would also provide something to show to potential investors. Bo agrees, we definitely want to build something, no doubt, but the way we build it must be collaborative since nobody has funds and expertise to do this alone.

Chris: one of the very attractive thing of BONSAI is that everybody can see the specific task that would benefit them and BONSAI without buying the all picture. He is for example interested in decentralized data and this may be a synergy with BONSAI while also supporting his own project requirements.

Tomas and Romain offered to work on Rest API for automated data harvesting routines. Below there is a complete list of follow ups to the meeting (proposed and potential).

Some links that were shared during the meeting:

- Industrial Ecology Dashboard: <https://github.com/IndEcol/Dashboard>
- Open energy modelling initiative: <http://www.openmod-initiative.org/>
  - And their unfinanced in-person meetings, every 6 months: <https://wiki.openmod-initiative.org/wiki/Events>
- SETAC North-America and their Roadmaps (concept taken from the electronics (PWB?) sector):
  - Inventory Model Description and Revision (ongoing): [https://c.ymcdn.com/sites/www.setac.org/resource/resmgr/advisory\\_groups\\_north\\_america/SNA\\_LCA\\_IG\\_Roadmap-Inventory.pdf?hhSearchTerms=](https://c.ymcdn.com/sites/www.setac.org/resource/resmgr/advisory_groups_north_america/SNA_LCA_IG_Roadmap-Inventory.pdf?hhSearchTerms=)
  - Uncertainty (ongoing, link missing)
  - Supporting Decision Makers with Life Cycle Assessment (finished 2014): [http://www.lifecycleinitiative.org/wp-content/uploads/2012/12/Roadmap-item\\_Supporting-Decision-Makers-with-LCA\\_Final-draft-for-comment.pdf](http://www.lifecycleinitiative.org/wp-content/uploads/2012/12/Roadmap-item_Supporting-Decision-Makers-with-LCA_Final-draft-for-comment.pdf)
- MFA community and possibility to *upload* data: <http://www.stan2web.net/>
- openLCA Nexus for downloading LCI data/databases in zolca format: <https://nexus.openlca.org/databases>

### **3. Discussing opportunity to suggest BONSAI as an add-on to UNEP's GLAD project (15 min)**

Andreas mentioned the work that the GLAD project is doing and in fact this point on the agenda aims at discussing potential synergies. Bo explained that the Life Cycle Initiative has recently launched a new phase where they are asking for people to contribute with projects ideas. If we look at GLAD from our community perspective is not really fulfilling our goal since theirs is mainly to make different databases talks together. But they may be interested in our objective as an add on to their initiative. In that case we could ask to make our work be part of the LC Initiative.

Andreas: it is still not clear how the new phase is going to work but there are supposed to be some more funds.

Mark: he was in the design of the new phase. There are radical changes since the new phase is more transparent and there are nw democratic elections before one had to pay to vote). The idea is to feed governments with information. This would also be the main source of incomes GLAD would be able to explain how this info may help governments. Thus he is supportive of trying but warn to be careful at the condition to comply to when you are awarded funds. Freedom may be more limited.

We will try to follow up on this in the next stage.

#### 4. Re-launching our current communication channels ([GitHub](#), [LinkedIn](#), [Working Groups](#), Board)

There is generally wish to keep in touch and follow up to this meeting. It may be a good idea to look at the wiki and the notes below first of all. Matchmaking received some consensus but requires funds for a coordinator. A periodical email summarizing activities and updates was also encouraged. The email/newsletter may not only focus on what happened within the BONSAI network but also outside e.g. GLAD, publications on related and relevant topics, since it is hard to keep up with all the news. It was proposed to have future meetings on specific focus point.

Expanding on the wiki is also a good idea, to document and keep track of future meetings, activities and developments.

Romain and Stefano proposed a mailing list for asking quick questions to the developers community.

Tair: on the GitHub there is a tab for hosting projects and address issues. This are two very useful tools.

Tomas: LinkedIn is already a good communication channel but we should use it more often.

Mark suggested that coordination can be part of sponsor packages since funds is a major obstacle. We could start a commercial/marketing group to deal with this. It is great what we are trying but we need also to solve the funding problem.

Bo: let's try to suggest a date for an in-person workshop.

### FOLLOW UP

#### Proposals for keeping the network productive:

- The BONSAI wiki on GitHub is just a draft and is meant to be an evolving cooperative effort. Michele made a few editing to make this clearer. The wiki is simply a coordinating tool at the service of the network. Clearly, the content may radically change as a consequence of future meetings and a consensus building phase. However, we seem to generally agree to use the BONSAI wiki on GitHub for this purpose and start sharing information of the kind:
  - Where your already developed software fits in the overall framework?
  - What are the missing pieces of code for a working beta version of the open access database?
  - What are the priority to address that currently are hindering progresses?  
(to be granted rights for editing the wiki email *michele.derosa@bonsai.uno*)
- We (Michele) will try to have a bi-monthly (every two months) email summarizing recent activities and initiatives. The news may be circulated on BONSAI social media too.
- Developers could also have bi-monthly *bonsai online* technical meetings following the circulation of the email with updates. Meetings could cover a specific topic proposed by the developers community. Propose topics in the table below.
- We are reflecting on possibility to organize *in-person* meeting/workshops for actual software development, e.g. in combination with major conferences, once or twice a year (**fill in Table 1 below for identify topic and priority**).

- We could have a first in-person meeting network building and identify immediate steps to take for allowing developments and attract investments. Mark proposed before or after [CSS](#) in Barcelona or to offer Pré rooms for a later date. If this is not possible we could have it online (**fill in table 2 below**).
- Marie de Saxcé (2.-0 LCA consultants) will connect with LCA developers working on ontologies to review the work which was done against existing glossaries ([ILCA glossary](#), [BONSAI glossary](#), [UNEP SETAC glossary](#) etc.).
- Romain and Tomas working on REST API for automating procedures of web scraping and data harvesting
- The table below is a sort of survey investigating priorities and potential topics for a first technical meeting. Please contribute by proposing a topic or supporting the proposed ones.

**Table 1** - proposed topic for first developers meeting and interested folk

<b>Topic for next technical online or in-person (developers) meeting</b>	<b>Proposed by</b>	<b>Interested folk</b>
<i>Data harmonization in a useful framework</i>	Stefan	Romain, Michele
<i>A common ontology and harmonized classification for Industrial Ecology</i>	Marie	Chris M., Bo, Michele
<i>Scientific article on the technology roadmap of BONSAI</i>	Bo	Chris M., Michele
<i>Interface to share inventory data</i>	Stefan	Romain, Michele
<i>Guidelines for data export</i>	Konstantin	
<i>Develop fundraising strategy</i>	Mark	Michele

**Table 2** - possible dates for a first in-person meeting with following draft agenda below:

Phase 1 -network building and following deliverables could be:

1. Updated vision of the immediate and final goal of BONSAI
2. Defining a governance structure, to identify priorities and a strategy to identify funds and how to get them
3. A plan for phase 2 in which we will combine the low hanging fruit in all the combined research and develop a proof of concept to show the collaboration can deliver something. We could also set up some specific working groups, where people focus on some key issues
4. A budget estimate for phase 2 (or perhaps some alternatives in case we have less funding for this phase than expected)

Possible next meeting date	Name of available people
Before LCA CSS Barcelona 24-25* Nov <b>(in person)</b>	
After LCA CSS Barcelona 29-30 Nov <b>(in person)</b>	
Early December <b>(online meeting)</b>	
A date to define in Amsterdam (Mark offers hosting space) <b>(in person)</b>	
Other suggestion? (...)	

\*Sunday already booked by LCA EU group discussion on ISO amendments