

Agenda:

- Alignment on the definition of what is intended for 'data integration' within the BONSAI framework
- Status on data integration and LCA, what do we know already?
- Identification of main gaps/problems for IE data integration
- Identification of potential working area
- Discussing next step
- Planning an in-person meeting
- AOB

MINUTES

Present: Michele De Rosa, Bo Weidema, Chris Mutel, Andreas Ciroth, Tomas Navarrete, Brandon Kuczenski, Pascal Lesage, Romain Sacchi, Stefano Merciai and Cyril Francois.

Absent but received the minutes: Ben Zhu, Blane Grann, Chris Davis, Guillaume Majeau-Bettez, Jannick Schmidt, Joke Dufourmont, Konstantin Stadler, Massimo Pizzol, Manuel Klarmann, Marie de Saxcé, Pascal Lesage, Radek Lonka, Stefan Pauliuk, Tair Bilyalov, Tom Gloria, Mark Goedkoop.

Meeting starts at 4:05.

The Bonsai wiki has a page dedicated to [Data Integration](#). Check it and integrate it.

Bo explained the idea of a living, i.e. continuously developing classification, where any new addition would either be a disaggregation of an existing class or an aggregation of existing classes. An infinite number of alternative aggregations is possible, thus making up different classifications.

Bo explained that classifications were first of all needed in the three dimensions:

- Flow-objects, i.e. intermediate flows (products, wastes), elementary flows, and flows in the environment (between environmental mechanisms)
- Activities, i.e. human production and consumption activities and environmental mechanisms
- Properties, which would include also properties that can be used to distinguish e.g. "high stack" and "rural"

Andreas suggested that there were an alternative option to make a "best classification", i.e. a new common-base nomenclature; a "central" classification to which all other existing ones are mapped. We should check the limitation of existing classification and avoid repeating errors. However, we are not aware of existing literature on challenges in mapping and/or creating classifications in IE.

Bo suggested that the "best classification" is always the most detailed one, because all other classifications can be mapped as aggregates of that. The living classification will at the same time provide correspondence tables/mapping between existing nomenclatures.

It was decided that the classifications should be hosted and public accessible on the BONSAI GitHub. We can start with whatever classifications people are currently willing to upload and map, leaving the possibility to evolve gradually.

There was a consensus about beginning to put on the BONSAI GitHub two existing classifications side by side (e.g. ISIC and NACE; http://ec.europa.eu/eurostat/statistics-explained/index.php/NACE_background) to allow merging them keeping the most disaggregated items, and then continuing with the one that becomes the most detailed classification as “*our continuously evolving classification*”. We could then do the same with other common LCA classifications. A suitable format could be JSON-LD. Each element should be tagged with the name of the classification it belongs to and a timestamp.

ISIC exists in excel too. Cyril has the classification in RDF format (<https://github.com/cfrancois7/IEO-ontology> - documentation will follow in a week or so), should be feasible to transform it in JSON. Andreas will make available the data format converter developed by the GLAD nomenclature group (ecoinvent and ILCD and the Japanese JSIC). It definitely needs further development but it is a starting point.

It would be fine to allow people to dump their classification in other formats, but it would only be integrated in the living classification in the moment that someone actually converts it to JSON-LD and solves any alignment issues with the currently most detailed classification.

Thomas suggested that there would also be a need for management software to help data providers and users to navigate, query and edit the classifications.

Action points before next meeting:

- Cyril: share on BONSAI wiki the ISIC classification in RDF format;
- Andreas: share on BONSAI wiki the data format converter developed by the GLAD nomenclature group (ecoinvent and ILCD and the Japanese JSIC)

Quick update on CDLCI initiative: At the upcoming meeting on the 26th partners shall provide:

- 1) Available data archetypes that can be delivered, in order to further discuss the technical concept / core more in detail;
- 2) Conditions for each partner to participate and the potential showstoppers.

The meeting ended at 5:10 pm CET.

Michele De Rosa and Bo Weidema