GreenDelta

sustainability consulting + software

JSON-LD: A smarter format for LCA data interchange

Andreas Ciroth, Michael Srocka

GreenDelta GmbH

New Delhi, ILCM 2015, September 2015

JSON-LD – Agenda

1. LCA data formats, status
2. Why a new, why yet another, data format
3. JSON-LD: principles
4. JSON-LD implementation status in LCA
5. Discussion, outlook and an invitation

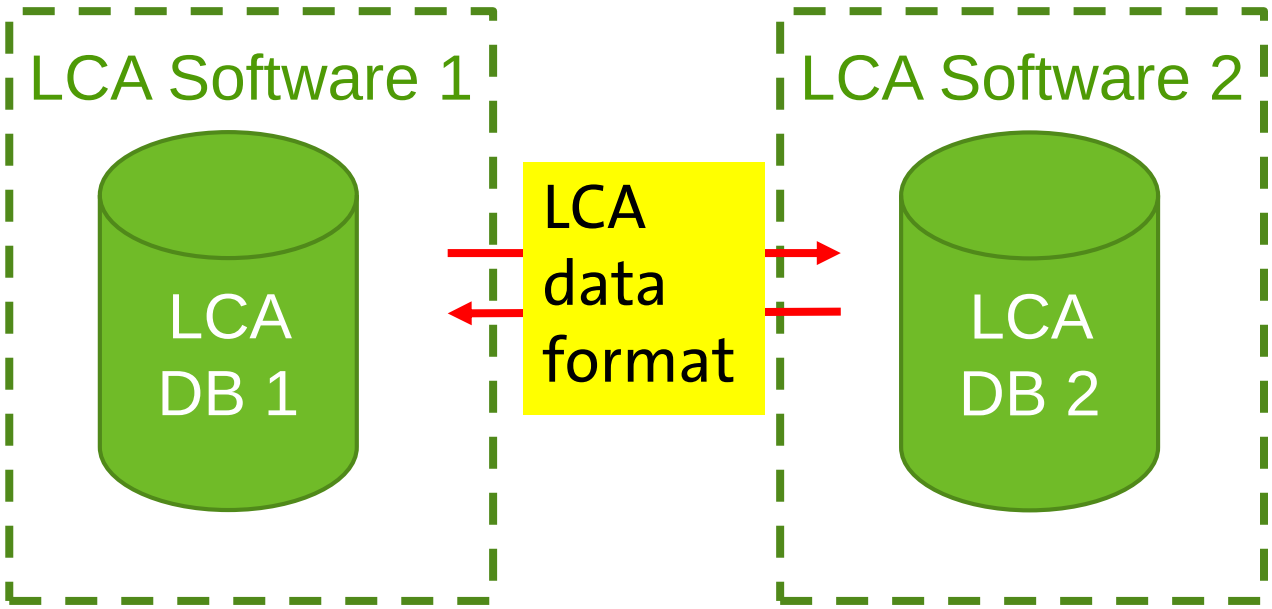
1 LCA data formats, status

LCA data formats, status

- Basically, a data format is simply a way and an agreement how to represent information.
- For LCA, data exchange is important, and therefore agreement on LCA data exchange formats is important as well.
- What we know as LCA data format nowadays are always LCA data exchange formats.

LCA data formats, status

- What we know as LCA data format nowadays are always LCA data exchange formats. The format which is used to store the data sets in a database and software is not directly relevant (!)



LCA data formats, status

- LCA data formats have a rather long history in LCA
- Spold99 was the first format to address the data exchange aspect (to my knowledge)
- Today, major formats are
  - EcoSpold1 (ecoinvent 2 database, all major LCA software systems, albeit in “dialects”)
  - ILCD (ELCD database of the JRC)
  - EcoSpold2 (ecoinvent 3 database)

EcoSpold1, EcoSpold2, ILCD format

- All are compliant with ISO14048/TS
- All are based on XML

```
<?xml version="1.0" encoding="utf-8"?>
<?xml-stylesheet version="1.0" href="../../stylesheets/process2html.xsl" type="text/xsl"?>
<processDataSet xmlns="http://lca.jrc.it/ILCD/Process" xmlns:common="http://lca.jrc.it/ILCD/Common" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://lca.jrc.it/ILCD/Process ../../schemas/ILCD_ProcessDataSet.xsd" locations="../ILCDLocations.xml" version="1.1">
  <processInformation>
    <dataSetInformation>
      <common:UUID>1d18b022-468b-45af-8b99-9fb095a33cc0</common:UUID>
      <name>
        <baseName xml:lang="en">Process steam from light fuel oil</baseName>
        <treatmentStandardsRoutes xml:lang="en">heat plant</treatmentStandardsRoutes>
        <mixAndLocationTypes xml:lang="en">consumption mix, at plant</mixAndLocationTypes>
        <functionalUnitFlowProperties xml:lang="en">MJ</functionalUnitFlowProperties>
      </name>
      <classificationInformation>
        <common:classification>
          <common:class level="0">Energy carriers and technologies</common:class>
          <common:class level="1">Heat and steam</common:class>
        </common:classification>
      </classificationInformation>
      <common:generalComment xml:lang="en">Good overall data quality. Energy carrier mix information based on official statistical information including import / export. A detailed
        heat plant model was used, which combine measured emissions plus calculated values for not measured emissions of e.g. organics or heavy metals. Energy carrier extraction
        and processing data is of sufficient to good (e.g. refinery) data quality. Inventory is partly based on primary industry data, partly on secondary literature
        data.</common:generalComment>
    </dataSetInformation>
    <quantitativeReference type="Reference flow(s)">
      <referenceToReferenceFlow>61</referenceToReferenceFlow>
    </quantitativeReference>
    <time>
      <common:referenceYear>2002</common:referenceYear>
      <common:dataSetValidUntil>2010</common:dataSetValidUntil>
      <common:timeRepresentativenessDescription xml:lang="en">Annual average</common:timeRepresentativenessDescription>
    </time>
```

EcoSpold1, EcoSpold2, ILCD format

- Of course, differences:
  - EcoSpold 1 rather simple (no parameters, only two languages, process and product is more or less the same...)
  - ILCD, EcoSpold2: more recent, quite complicated, typically several thousand, interlinked files for one process data set
  - Unfortunately all not fully consistent with each other

EcoSpold1, EcoSpold2, ILCD format

- Unfortunately all not fully consistent with each other

→ openLCA format converter and openLCA import export features help to work with different formats but not a perfect solution

→ Effort in implementation
(e.g.: ILCD: every text field in different language possible)

→ Inconsistencies in formats
(e.g. flow properties: ILCD: one value for one flow; EcoSpold2: one value for one exchange, may be different for each process)

2 Why a new, why yet another, data format

Why a new format

In a nutshell:

- New technological developments in “resource description” promise to be better for LCA format specification
- A new, overarching format allows to consider the existing aspects of different existing formats better.

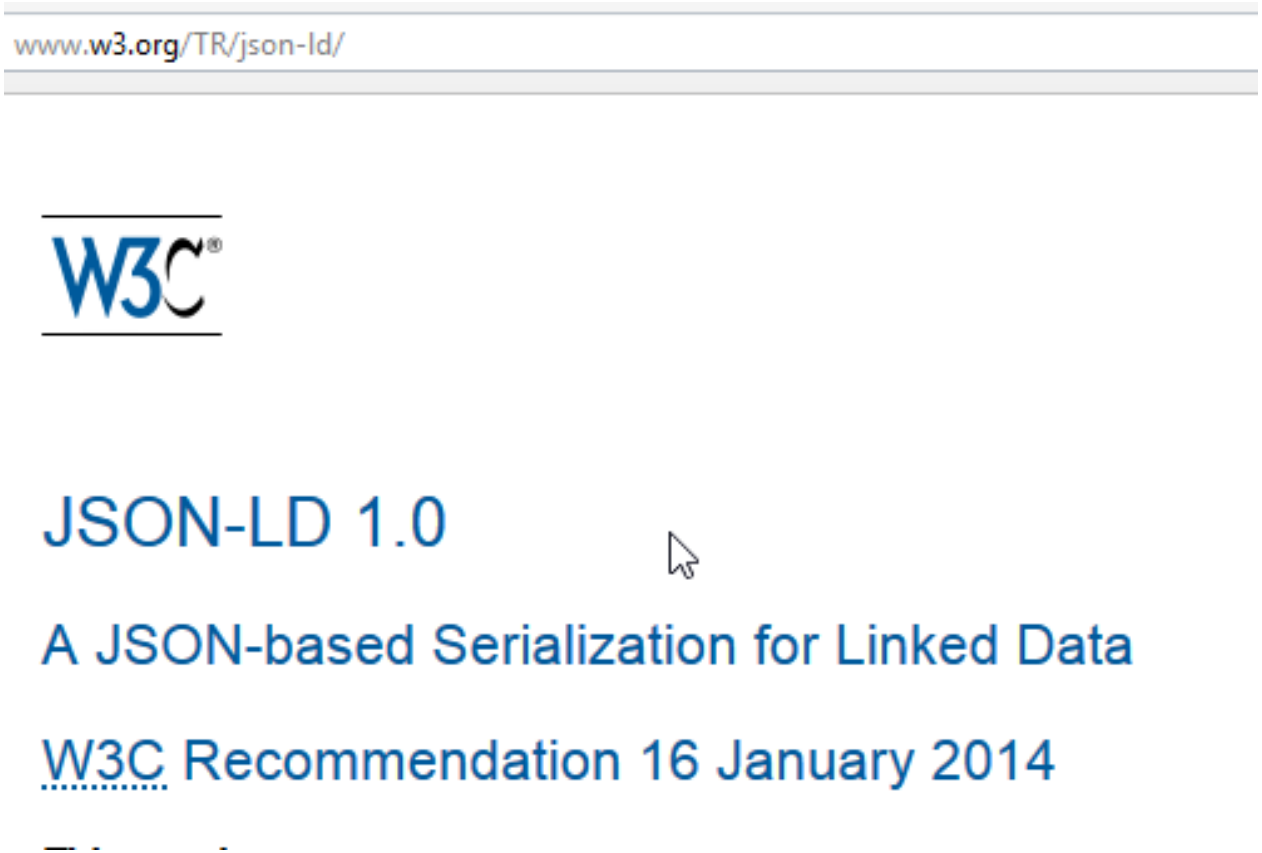
3 JSON-LD: principles

JSON-LD: javascript object notation for linked data
Specified by the W3C

- JSON-LD: “JavaScript Object Notation for Linked Data”, W3C recommended for linked data (14 January 2014)
- Can directly be parsed as RDF triples, therefore direct link to ontologies (RDF format)
- Lightweight, human-readable, Google, Yahoo, Yandex, … are supporters
- Based on JSON (json.org) – the data format of the web, parsers available, really easy integration..

JSON-LD: javascript object notation for linked data

Specified by the W3C



Starting point: schema.org

Product

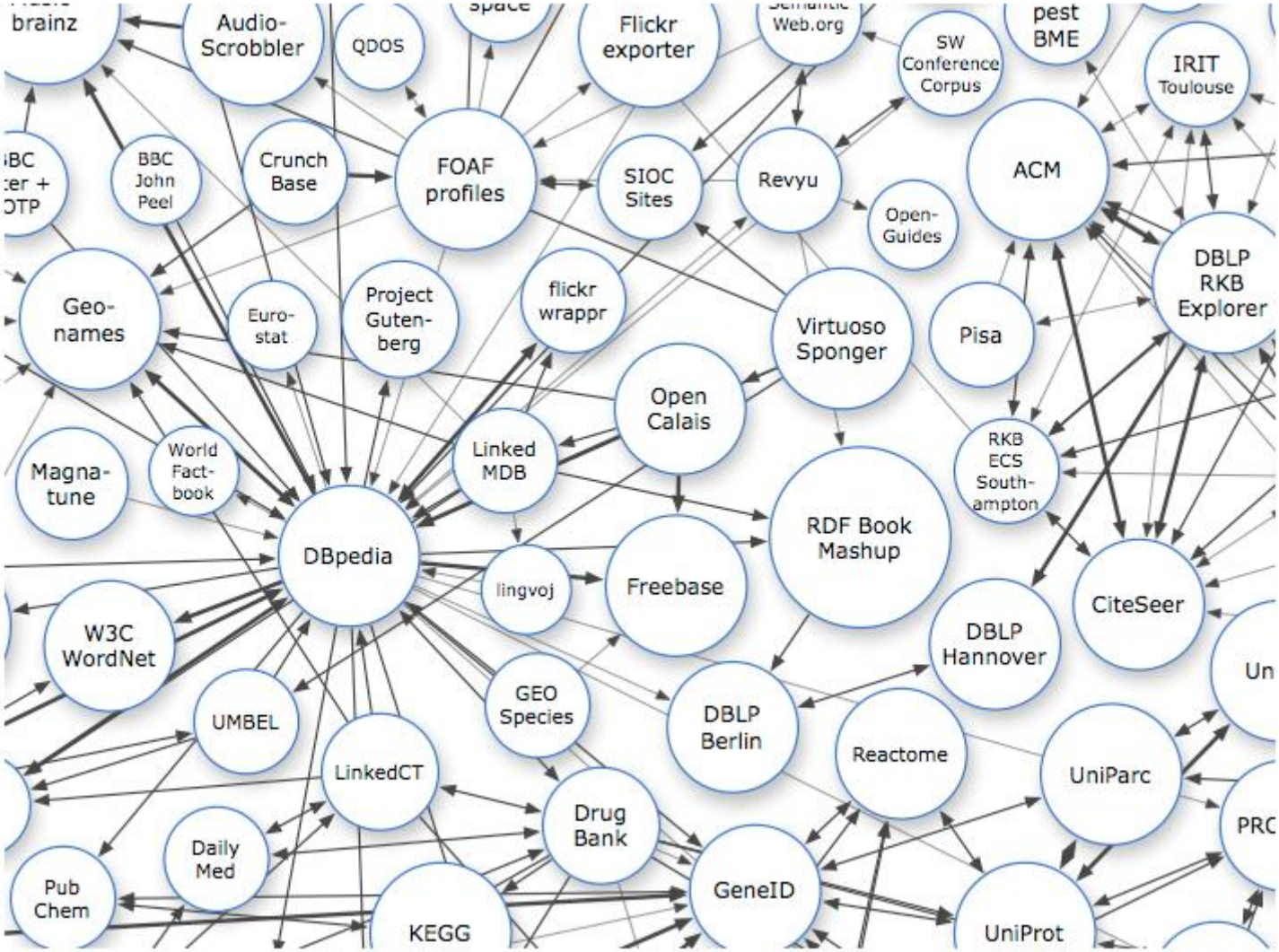
Thing > Product

Any offered product or service. For example: a pair of shoes; a concert ticket; the rental of a car; a haircut

Usage: Over 1,000,000 domains

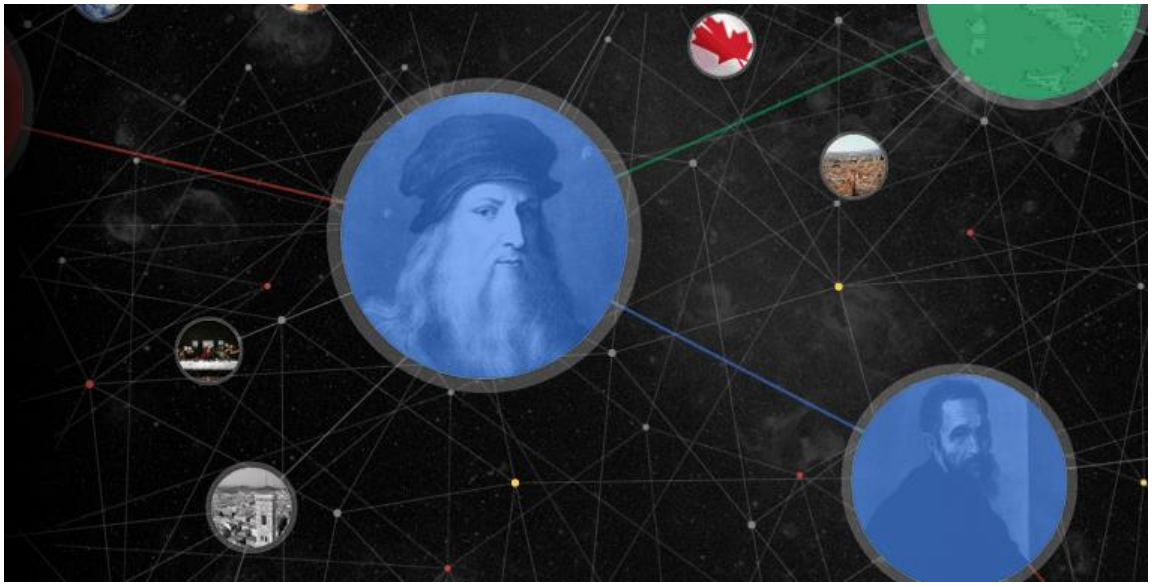
Property	Expected Type	Description
Properties from Product		
additionalProperty	PropertyValue	A property–value pair representing an additi or another characteristic for which there is n Note: Publishers should be aware that applic (e.g. http://schema.org/width, http://schem expect such data to be provided using those

Linked data


brainz
Audio-Scrobbler
QDOS
Flickr exporter
Web.org
SW Conference Corpus
pest BME
IRIT Toulouse
BBC John Peel
Crunch Base
FOAF profiles
SIOC Sites
Revyu
Open-Guides
ACM
DBLP RKB Explorer
Geo-names
Euro-stat
Project Guten-berg
flickr wrappr
Virtuoso Sponger
Pisa
Open Calais
Linked MDB
Magna-tune
World Fact-book
RKB ECS South-ampton
DBpedia
lingvoj
Freebase
RDF Book Mashup
CiteSeer
W3C WordNet
DBLP Hannover
UMBEL
GEO Species
DBLP Berlin
Reactome
UniParc
LinkedCT
Drug Bank
Daily Med
Pub Chem
KEGG
GeneID
UniProt

Knowledge Graphs

- Used by Google to enhance the search results;
- Content from several sources
- Basically objects and their relations
- Provide structured and detailed information about the topic in addition to a list of links to other sites
- A very practical and relevant implementation of linked data

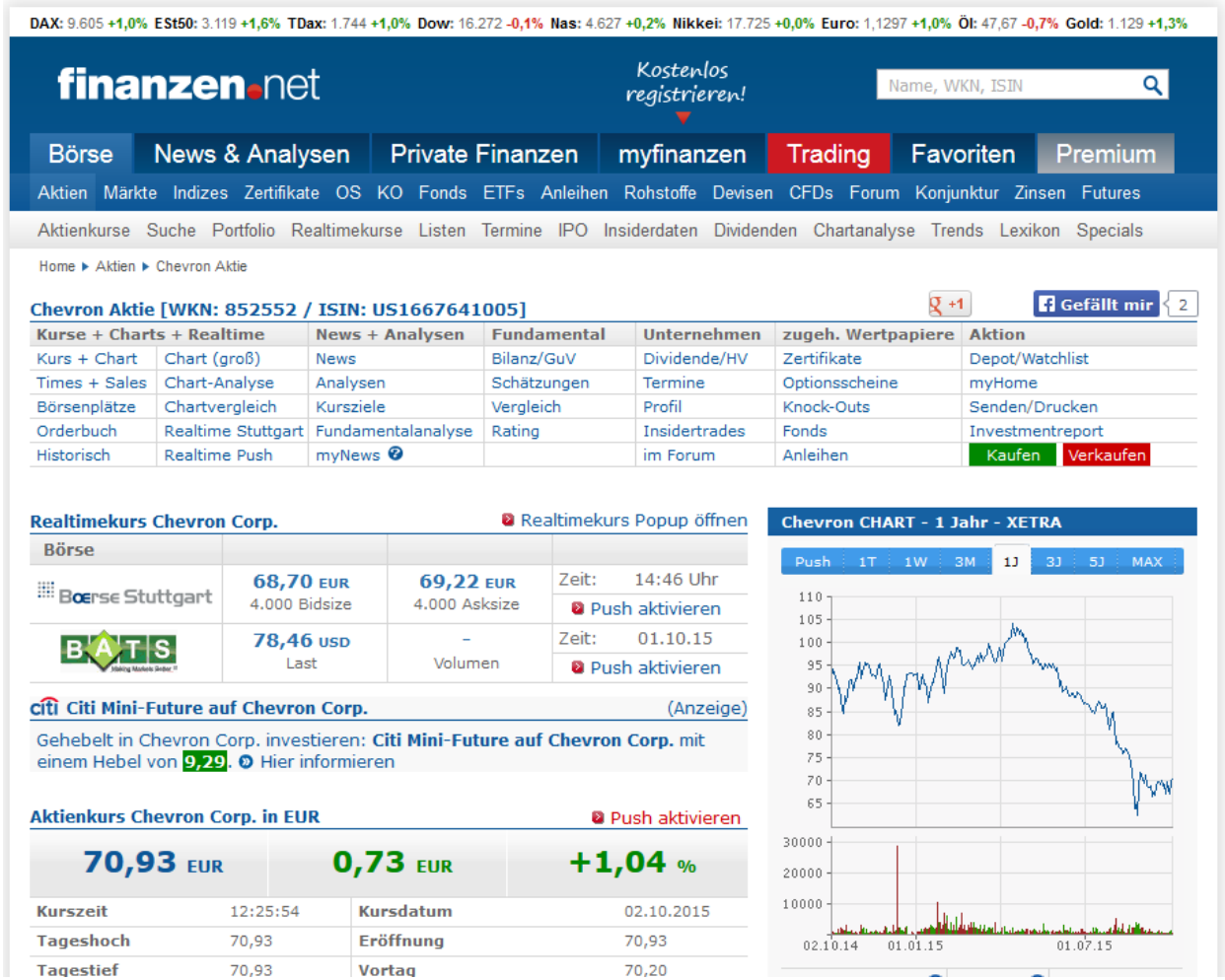


JSON-LD can be embedded in HTML

(i.e. in normal websites)

```
<script type="application/ld+json">
{
  "@context": "http://schema.org",
  "@type": "Person",
  "name": "John Doe",
  "jobTitle": "Graduate research assistant",
  "affiliation": "University of Dreams",
  "additionalName": "Johnny",
  "url": "http://www.example.com",
  "address": {
    "@type": "PostalAddress",
    "streetAddress": "1234 Peach Drive",
    "addressLocality": "Wonderland",
    "addressRegion": "Georgia"
  }
}
</script>
```

e.g., http://www.finanzen.net/aktien/Chevron-Aktie



DAX

mehr

Commerzbank

mehr

Apple

mehr

Facebook

mehr

e.g., http://www.finanzen.net/aktien/Chevron-Aktie

Information harvesting, e.g. Chevron

(in Google search result, information display)

Chevron Corporation Home - Human Energy
www.**chevron**.com/ ▾ Diese Seite übersetzen
Chevron works to meet the world's growing demand for energy by exploring for oil and natural gas; refining and marketing gasoline; producing chemicals and ...
Chevron Careers Home - Contact - Investor Relations - Chevron Worldwide

Chevron Corporation – Wikipedia
https://de.wikipedia.org/wiki/**Chevron**_Corporation ▾
Chevron Corporation ist ein weltweit operierender Energiekonzern. Er gehört zu den weltgrößten Ölkonzernen und nach eigenen Angaben ist **Chevron** der ...
Geschichte - Umweltpolitik und Regenwald - Blockierung der Batterien von ...

Chevron – Wikipedia
https://de.wikipedia.org/wiki/**Chevron** ▾
Chevron (Uniform), am Ärmel der Soldatenuniform aufgebrachte Rangabzeichen, Dienstaltersbezeichnungen, Laufbahnabzeichen oder Auszeichnungen ...

Chevron Corporation

Chevron Corporation ist ein weltweit operierender Energiekonzern. Er gehört zu den weltgrößten Ölkonzernen und nach eigenen Angaben ist Chevron der größte Produzent von geothermischer Energie. Wikipedia

Aktienkurs: CHV (FRA)
69,97 € +0,13 (+0,18 %)
2. Okt., 14:31 MESZ - Haftungsausschluss

CEO: John S. Watson

Hauptsitz: San Ramon, Kalifornien, Vereinigte Staaten

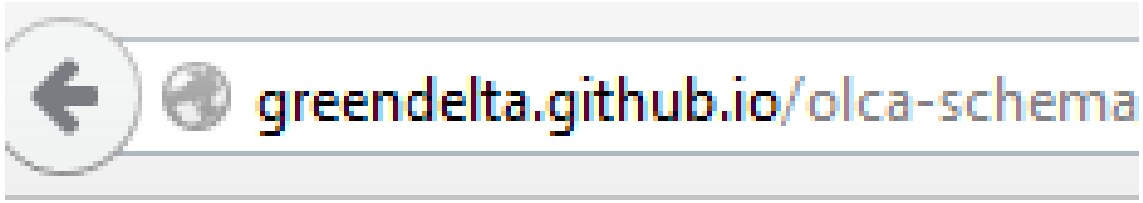
Gegründet: 10. September 1879, Kalifornien, Vereinigte Staaten

Feedback

4 JSON-LD implementation status in LCA

JSON-LD: implementation status in LCA

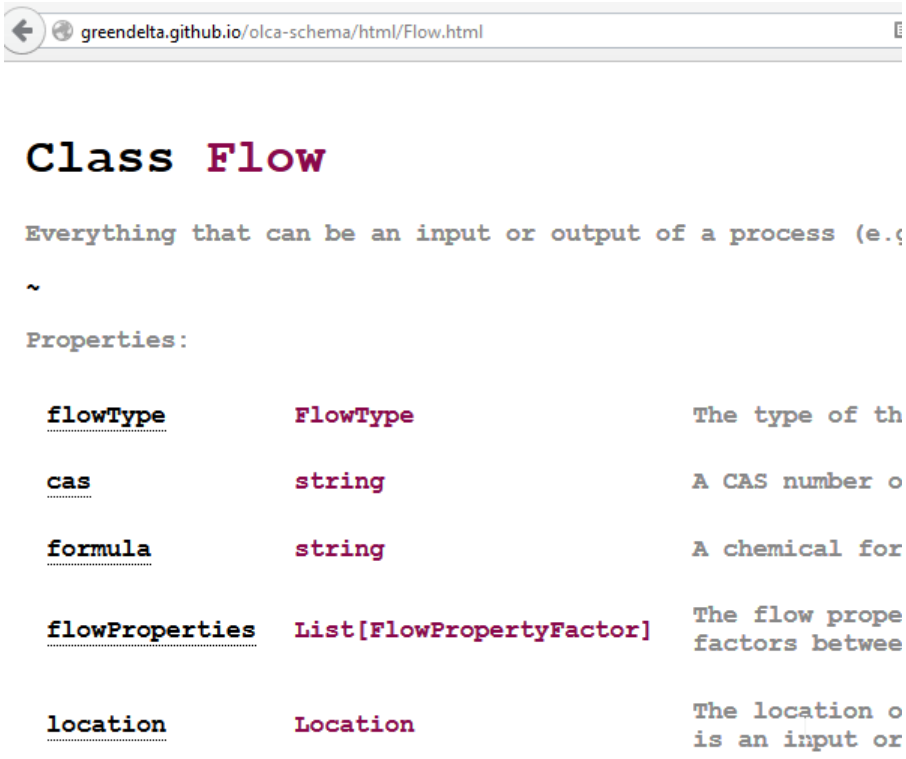
- The JSON-LD format is implemented as one of the formats in openLCA since summer 2015, based on a US EPA project
- The format implementation is openly documented, here:



- We follow the schema.org principles (display, documentation, ...)

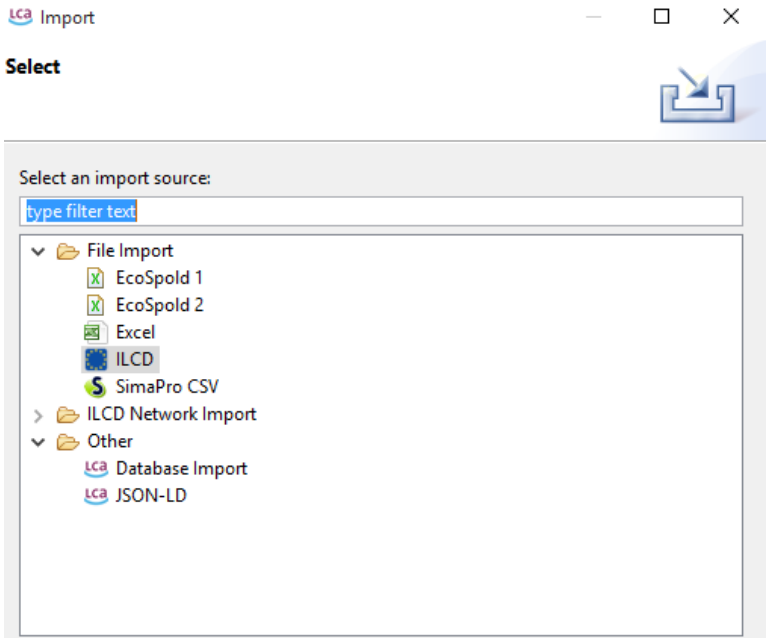
JSON-LD: implementation status in openLCA

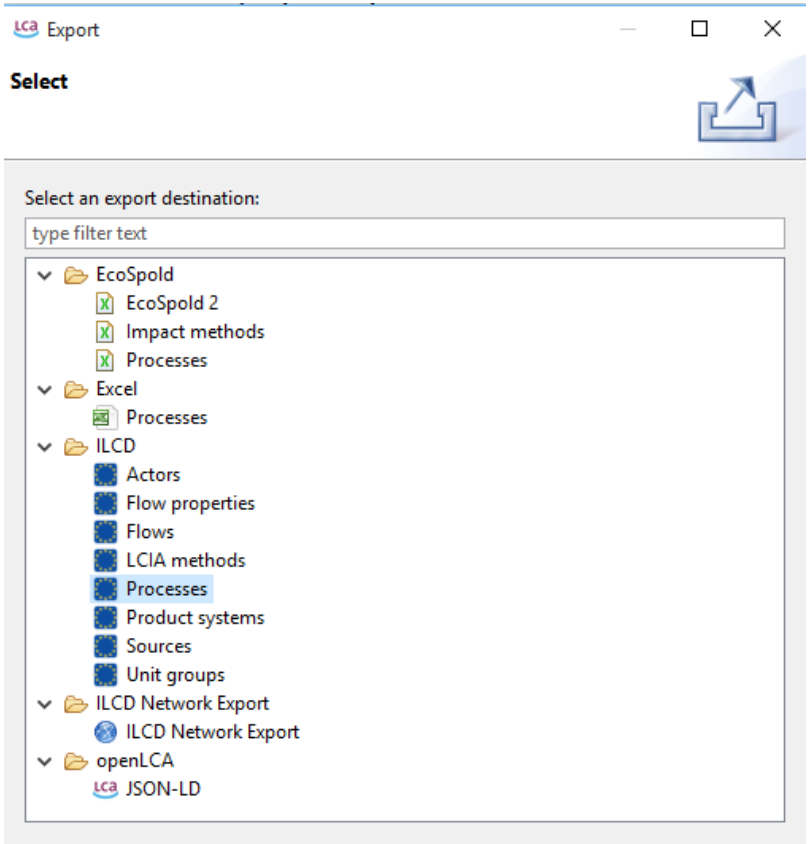
- A first version of the format fully implemented
- But of course interesting to put for discussion



JSON-LD: implementation status in openLCA

- The format is offered as additional format, other formats are of course still supported





ILCD format, process data set, one exchange:

disposal wood ash mixture, export from openLCA

```
<ns2:exchange xmlns:olca="http://openlca.org/ilcd-extensions" dataSetInternalID="50"
olca:unitId="20aadc24-a391-41cf-b340-3e4529f44bde" olca:propertyId=
"93a60a56-a3c8-11da-a746-0800200b9a66" olca:amount="1.1464E-4">
    <ns2:referenceToFlowDataSet type="flow data set" refObjectId=
    "02da3b87-8de6-37f5-be94-204c54304356" version="01.00.000" uri=
    "../flows/02da3b87-8de6-37f5-be94-204c54304356">
        <shortDescription xml:lang="en">disposal, wood ash mixture, pure, 0% water, to
        landfarming</shortDescription>
    </ns2:referenceToFlowDataSet>
    <ns2:exchangeDirection>Input</ns2:exchangeDirection>
    <ns2:meanAmount>1.1464E-4</ns2:meanAmount>
    <ns2:resultingAmount>1.1464E-4</ns2:resultingAmount>
    <ns2:uncertaintyDistributionType>log-normal</ns2:uncertaintyDistributionType>
    <ns2:relativeStandardDeviation95In>1.159784462734350096724256218294613063335418701171875
    </ns2:relativeStandardDeviation95In>
</ns2:exchange>
```

JSON-LD format, process data set, one exchange:

disposal wood ash mixture, export from openLCA


```
{
    "@type": "Exchange",
    "avoidedProduct": false,
    "input": true,
    "amount": 1.1464E-4,
    "flow": {
        "@type": "Flow",
        "@id": "d52f395a-a097-3cd5-b4d5-f0fb981745ec",
        "name": "disposal, wood ash mixture, pure, 0% water, to municipal incineration"
    },
    "unit": {
        "@type": "Unit",
        "@id": "20aadc24-a391-41cf-b340-3e4529f44bde",
        "name": "kg"
    },
    "flowProperty": {
        "@type": "FlowProperty",
        "@id": "93a60a56-a3c8-11da-a746-0800200b9a66",
        "name": "Mass"
    },
    "uncertainty": {
        "@type": "Uncertainty",
        "distributionType": "LOG_NORMAL_DISTRIBUTION",
        "geomMean": 1.1464E-4,
        "geomSd": 1.15978446273435
    }
```

5 Discussion, outlook and an invitation

Discussion

A new LCA data format should not be introduced without really good reasons.

We think there are good reasons to introduce JSON-LD as new format:

1. It is a very useful alternative to XML (i.e., EcoSpold, EcoSpold2, ILCD)
2. It helps overcome shortcomings of existing LCA data formats, and can fully replace them
3. It overcomes also shortcomings of RDF formats, more efficient and faster

Outlook

- openLCA implementation will be extended to LCIA models
- Linkage with the Life Cycle Harmonization Tool available (once this is released)
- (hopefully): Discussion, on an international level

..and an invitation:

- For other LCA tool developers: Please have a look, we of course think this format makes sense also in other tools; less implementation effort than other formats, much more powerful.
- Chance to explore this format in the context of the UNEP intergovernmental panel on interoperability

GreenDelta

sustainability consulting + software

Thank you

Contact: Dr. Andreas Ciroth

GreenDelta GmbH

Müllerstrasse 135, 13349 Berlin, Germany

ciroth@greendelta.com

www.greendelta.com