ETH zürich

Institute of Molecular Systems Biology

Where is the “*L*” in “ORD”?

Open Research Data – are we on track?
Hotel Kreuz, Zeughausgasse 41, Bern

Adriano Rutz, 2024-05-31

D BIOL

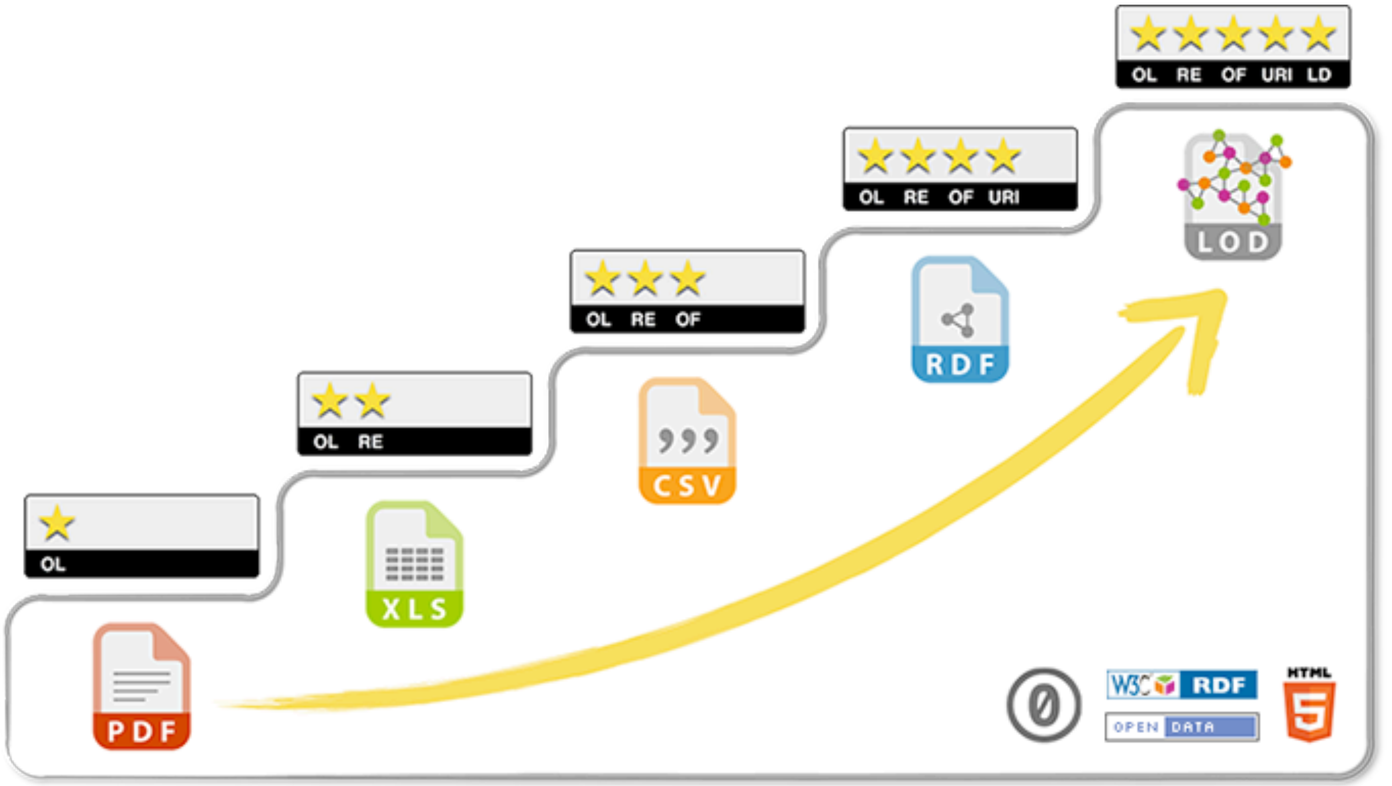
Unless otherwise indicated, content is licensed under CC-BY-SA 4.0

Disclaimer

The “*L*” stands for *linked*,
as in “Linked Open (Research) Data”

What is Linked Open Data (LOD)?

What is Linked Open Data (LOD)?

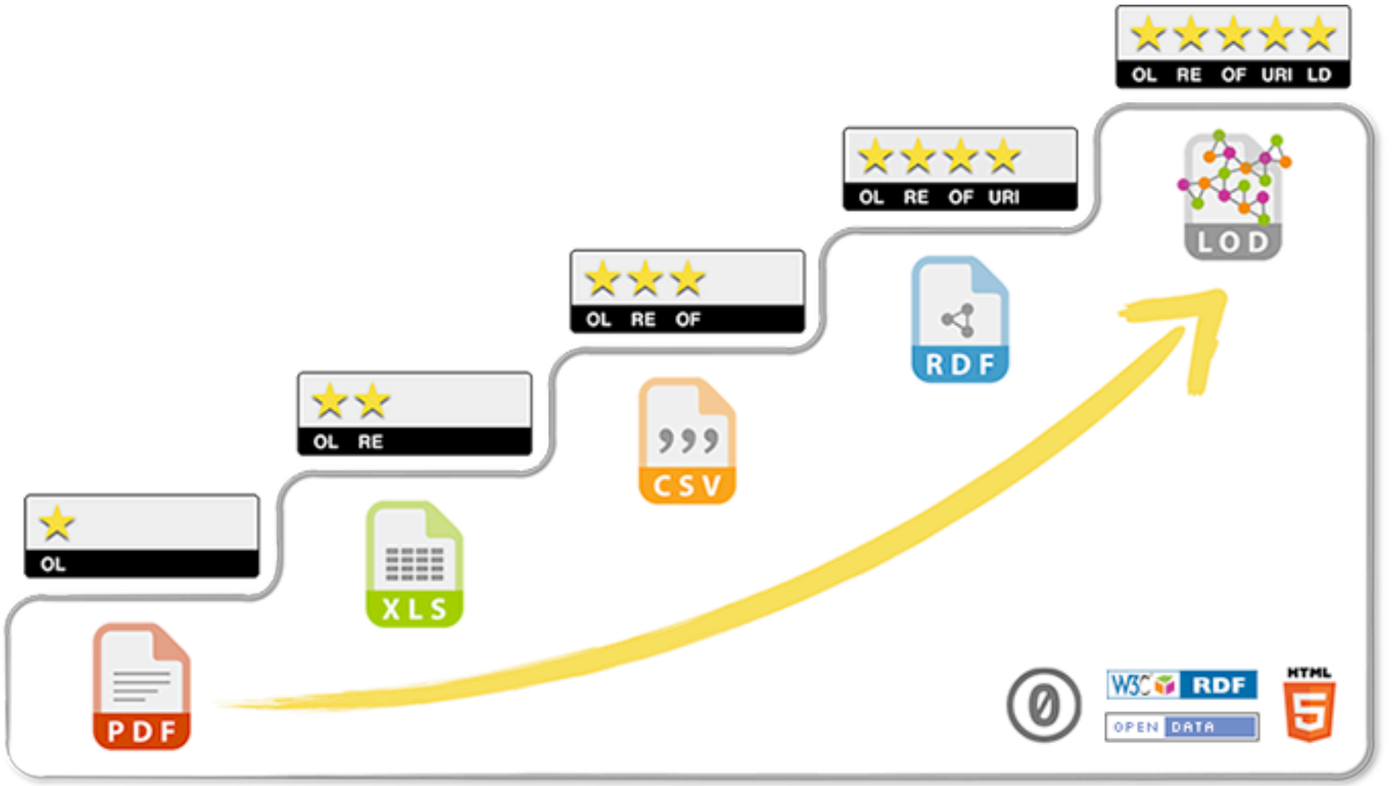


https://commons.wikimedia.org/wiki/File:5-star_deployment_scheme_for_Open_Data.png

https://5stardata.info

What is Linked Open Data (LOD)?

make your *content* available on the Web (whatever format) under an open license



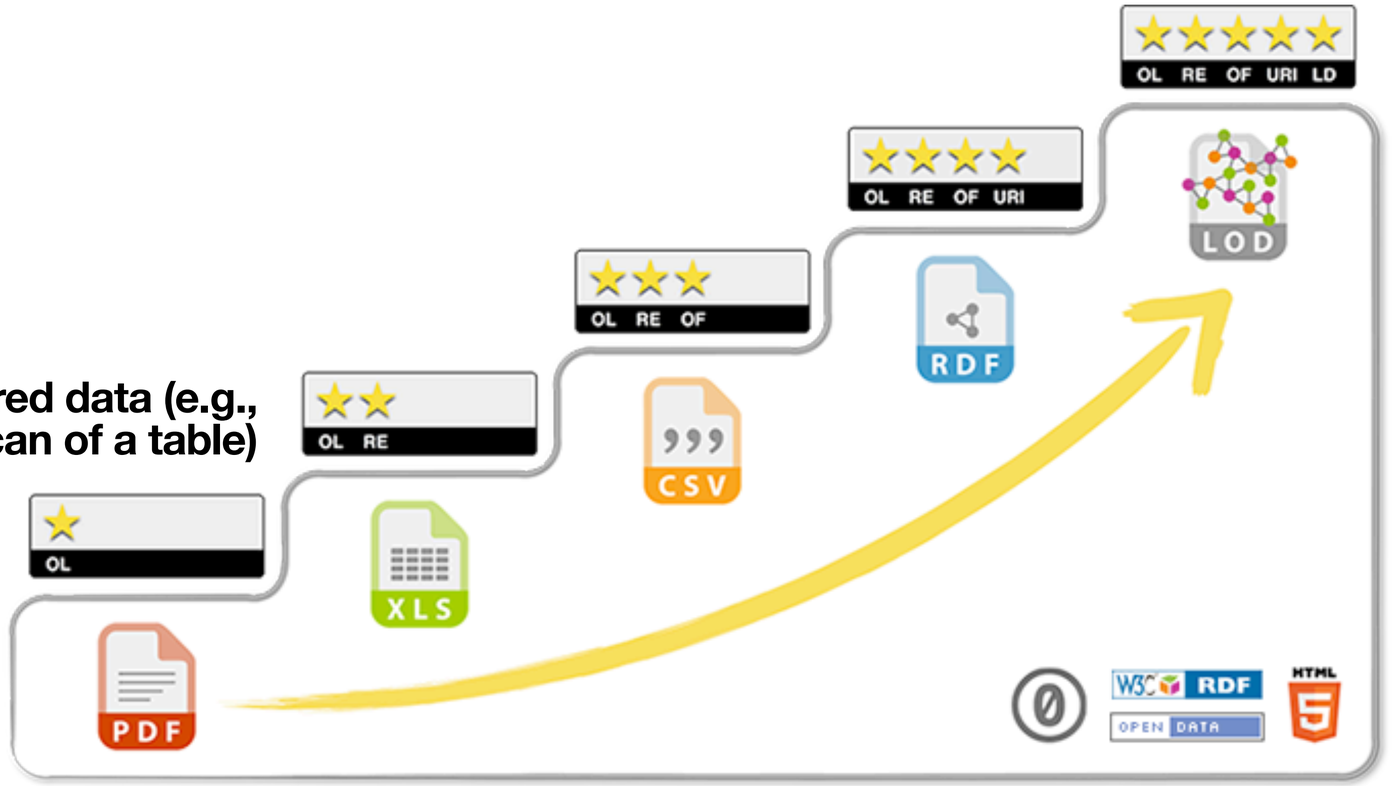
https://commons.wikimedia.org/wiki/File:5-star_deployment_scheme_for_Open_Data.png

https://5stardata.info

What is Linked Open Data (LOD)?

make it available as structured data (e.g., Excel instead of image scan of a table)

make your *content* available on the Web (whatever format) under an open license



https://commons.wikimedia.org/wiki/File:5-star_deployment_scheme_for_Open_Data.png

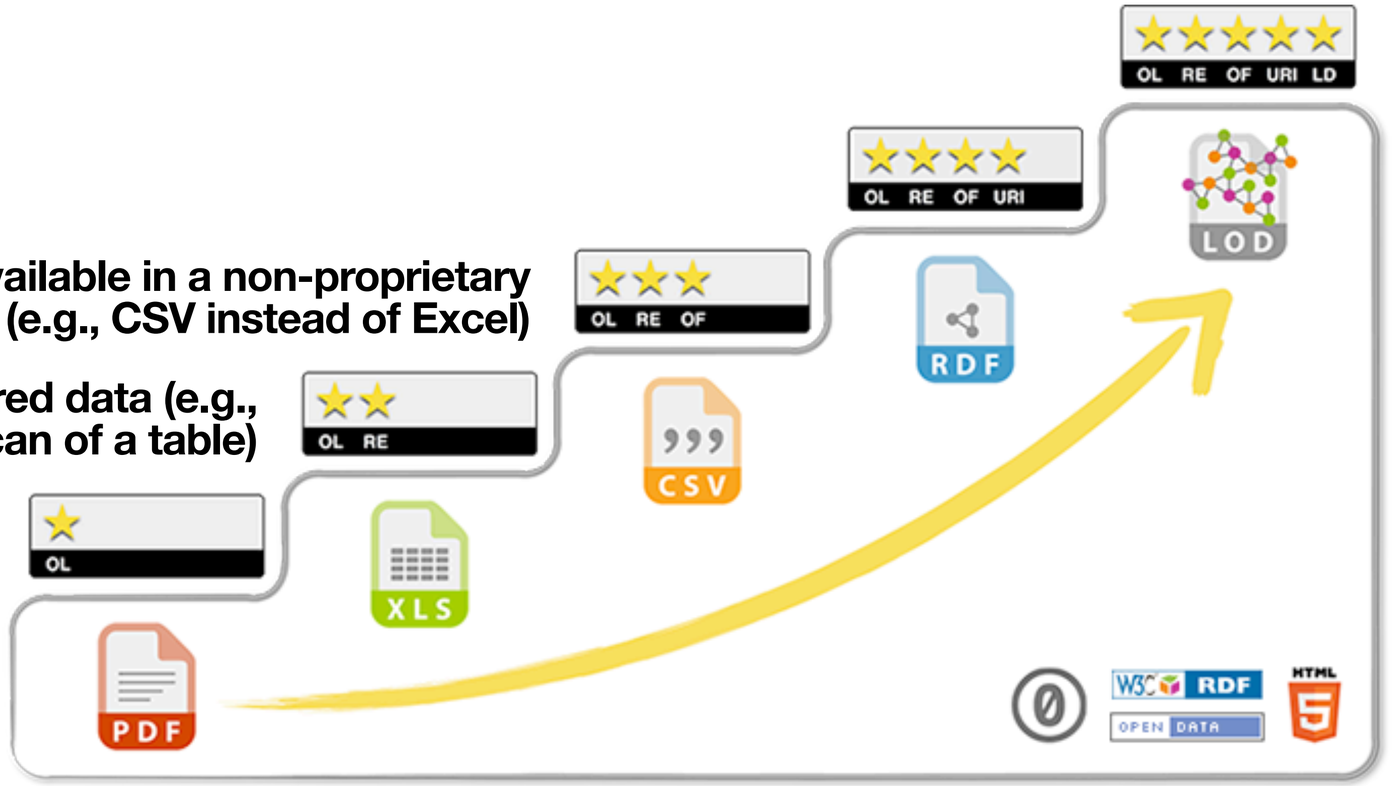
https://5stardata.info

What is Linked Open Data (LOD)?

make it available in a non-proprietary open format (e.g., CSV instead of Excel)

make it available as structured data (e.g., Excel instead of image scan of a table)

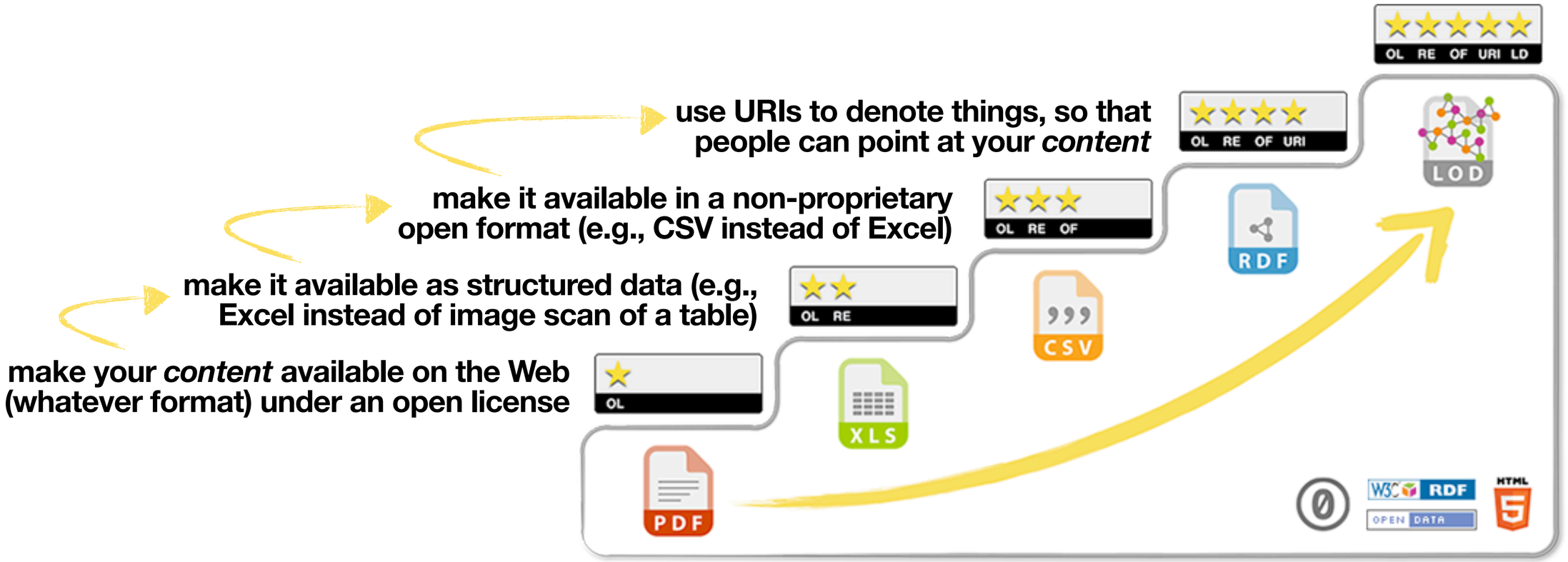
make your *content* available on the Web (whatever format) under an open license



https://commons.wikimedia.org/wiki/File:5-star_deployment_scheme_for_Open_Data.png

https://5stardata.info

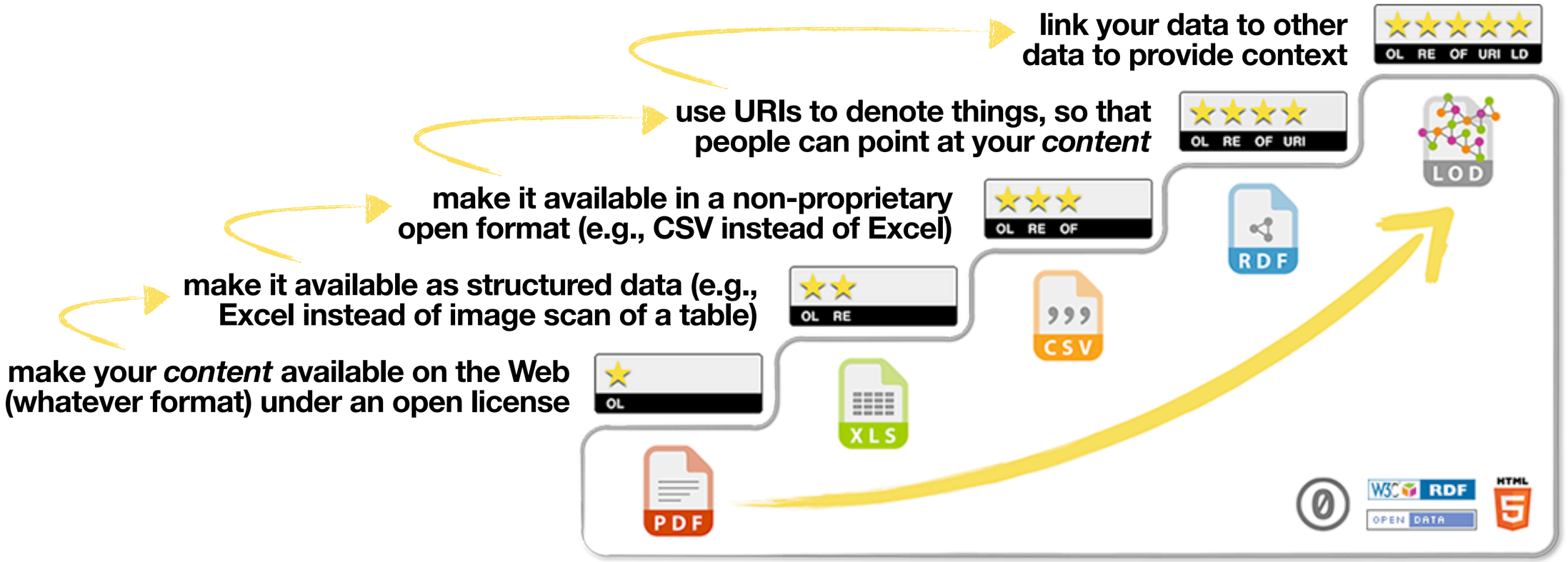
What is Linked Open Data (LOD)?



https://commons.wikimedia.org/wiki/File:5-star_deployment_scheme_for_Open_Data.png

https://5stardata.info

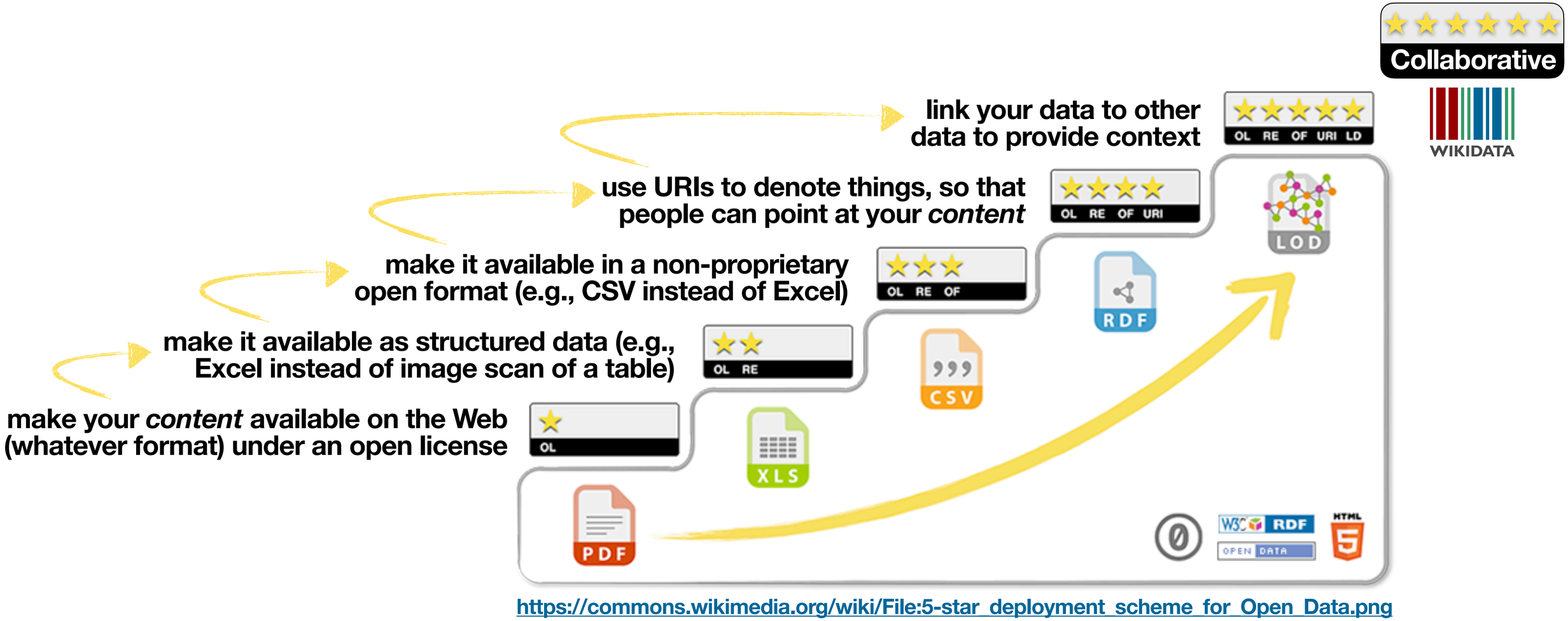
What is Linked Open Data (LOD)?



https://commons.wikimedia.org/wiki/File:5-star_deployment_scheme_for_Open_Data.png

https://5stardata.info

What is Linked Open Data (LOD)?



https://commons.wikimedia.org/wiki/File:5-star_deployment_scheme_for_Open_Data.png

https://5stardata.info

The “L”OTUS initiative

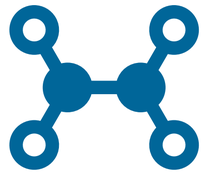
Available Natural Product Data

The “L”OTUS initiative

Available Natural Product Data

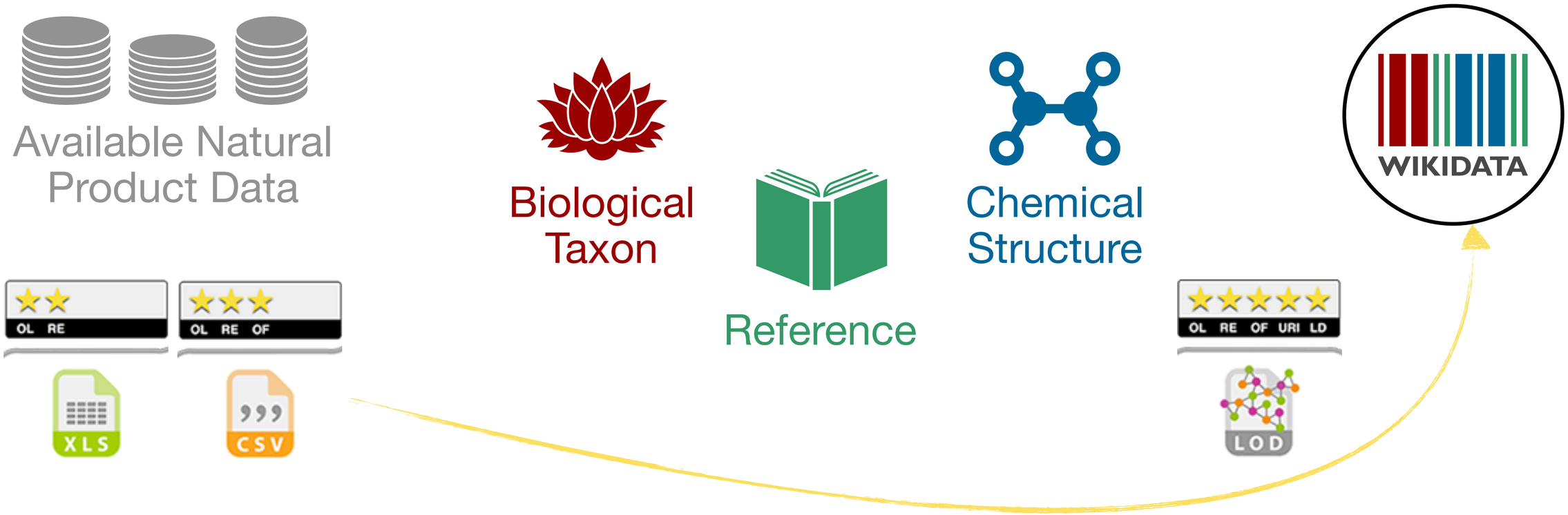
Biological Taxon

Reference



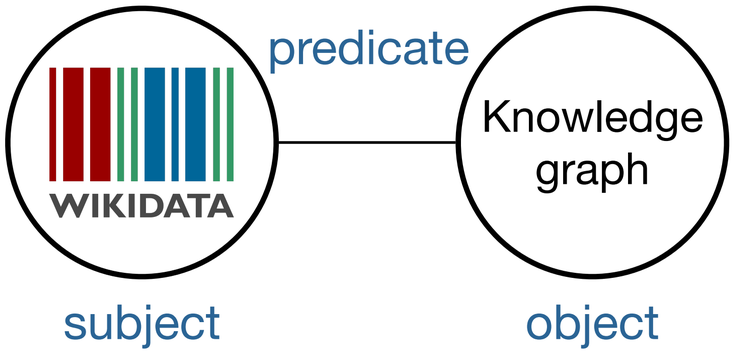
Chemical Structure

The “L”OTUS initiative



The “L”OTUS initiative - Wikidata

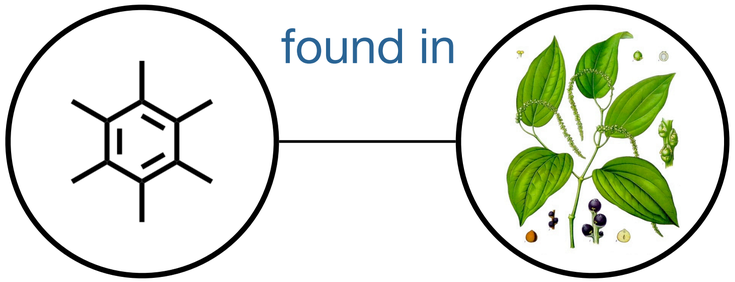
The “L”OTUS initiative - Wikidata



The “L”OTUS initiative - Wikidata

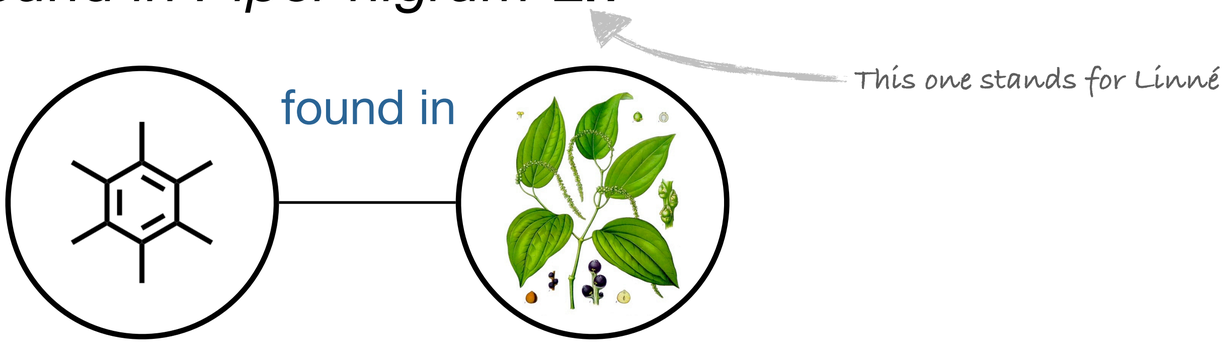
The initiative - SPARQ“L”

« Hey Wiki, what are the compounds found in Piper nigrum L.*? »*



The initiative - SPARQ“L”

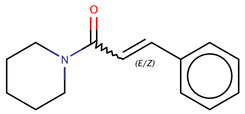
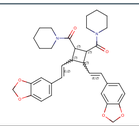
« Hey Wiki, what are the compounds found in Piper nigrum L.*? »*



The initiative - SPARQ“L”

« Hey Wiki, what are the compounds found in Piper nigrum L.*? »*

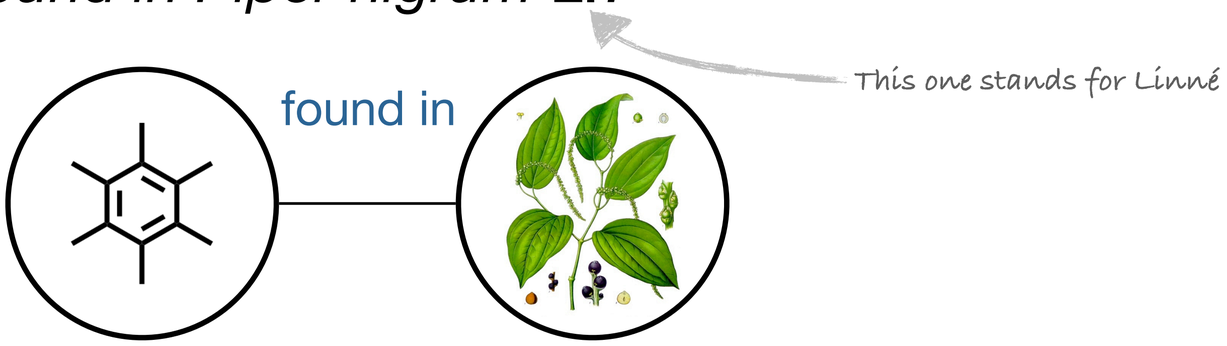
```
SELECT DISTINCT
?structure ?structure_smiles ?structure_inchikey
WHERE {
    VALUES ?organism {
        wd:Q43084 # Piper nigrum
        }
        ?organism_child (wdt:P171*) ?organism.
        ?structure wdt:P233 ?structure_smiles;
            wdt:P235 ?structure_inchikey;
            (p:P703/ps:P703) ?organism_child.
}
```

row	structure	structure_smiles	structure_inchikey
1	wd:Q105143500		KNOXUMZPTHELAO-UHFFFAOYSA-N
…	…	…	…
419	wd:Q105224845		QOESZPJYTVEXNN-UHFFFAOYSA-N

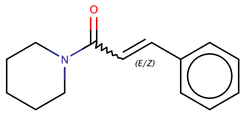
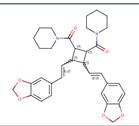
419 results in 0.5 second

The initiative - SPARQ“L”

« Hey Wiki, what are the compounds found in Piper nigrum L.*? »*



```
SELECT DISTINCT
?structure ?structure_smiles ?structure_inchikey
WHERE {
    VALUES ?organism {
        wd:Q43084 # Piper nigrum
        }
        ?organism_child (wdt:P171*) ?organism.
        ?structure wdt:P233 ?structure_smiles;
            wdt:P235 ?structure_inchikey;
            (p:P703/ps:P703) ?organism_child.
}
```

row	structure	structure_smiles	structure_inchikey
1	wd:Q105143500		KNOXUMZPTHELAO-UHFFFAOYSA-N
…	…	…	…
419	wd:Q105224845		QOESZPJYTVEXNN-UHFFFAOYSA-N

419 results in 0.5 second

The initiative - SPARQ“L”

« Hey Wiki, what are the compounds found in Gentianaceae? »

984 results in 0.9 second

The initiative - SPARQ“L”

« Hey Wiki, what are the compounds found in Gentianaceae? »

984 results in 0.9 second

« Hey Wiki, how many articles does that correspond to? »

396 results in 0.6 second

The initiative - SPARQ“L”

« Hey Wiki, what are the compounds found in Gentianaceae? »

984 results in 0.9 second

« Hey Wiki, how many articles does that correspond to? »

396 results in 0.6 second

« And how many of them are open access? »

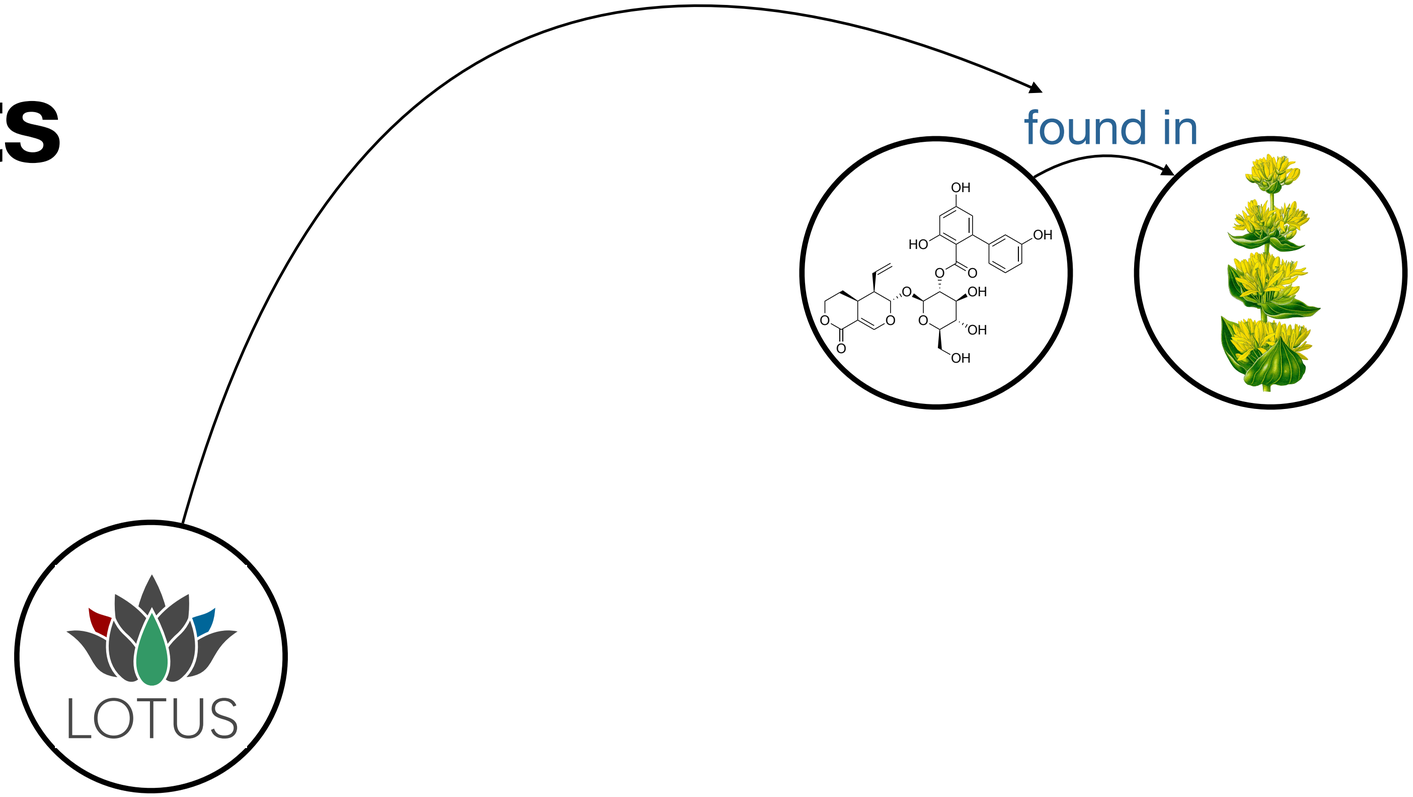
67 results in 0.6 second

The initiative - SPARQ“L”

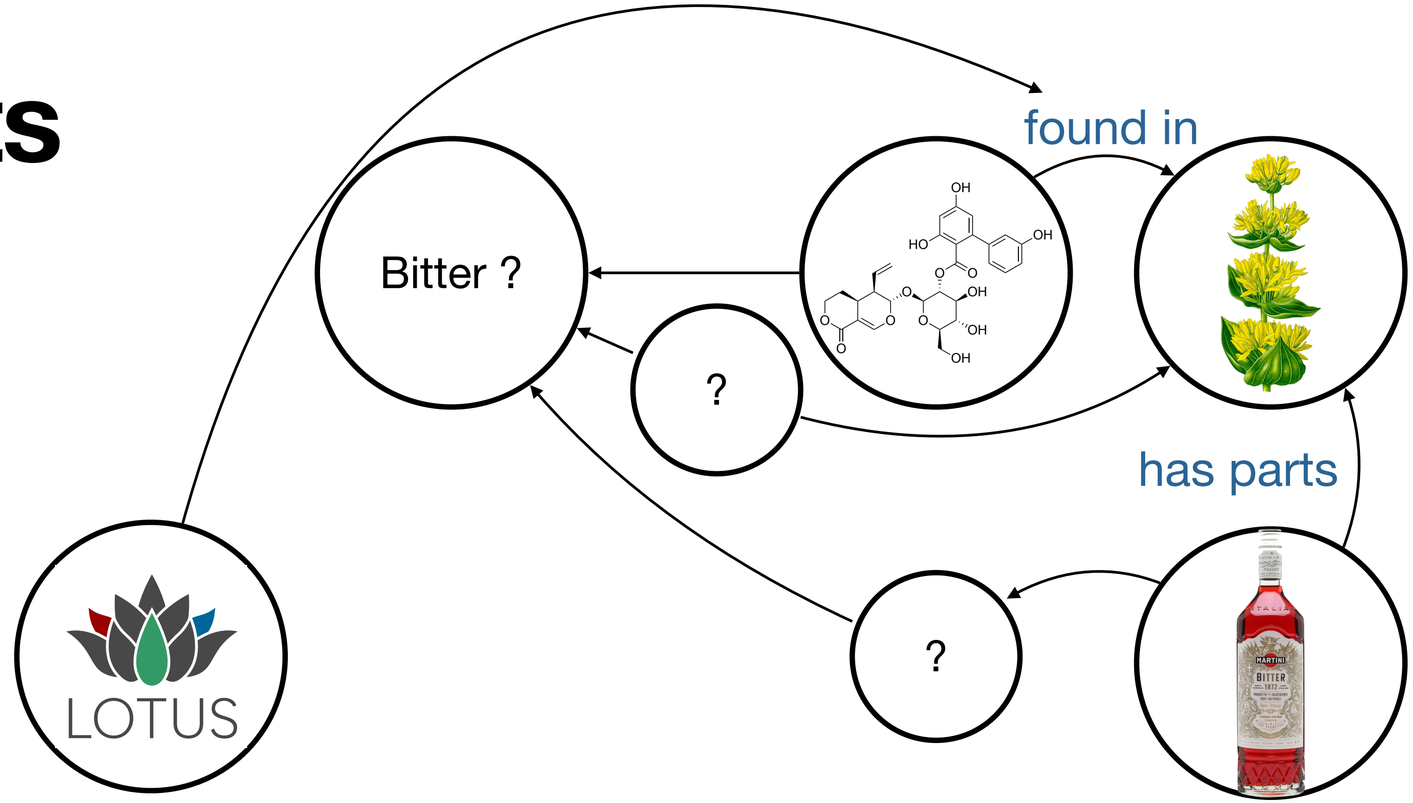
« And how many of them are open access? »

67 results in 0.6 second

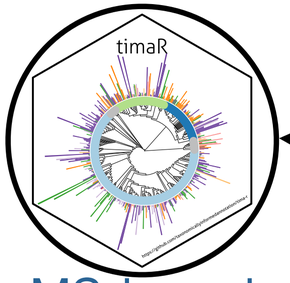
LO(R)D has no limits



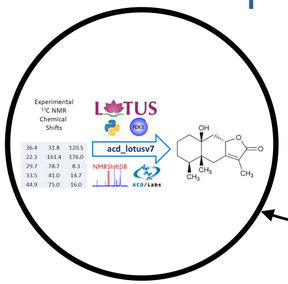
LO(R)D has no limits



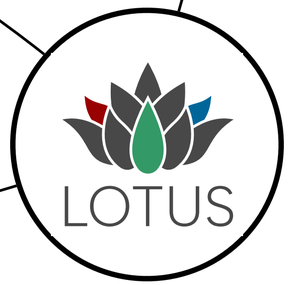
LO(R)D has no limits

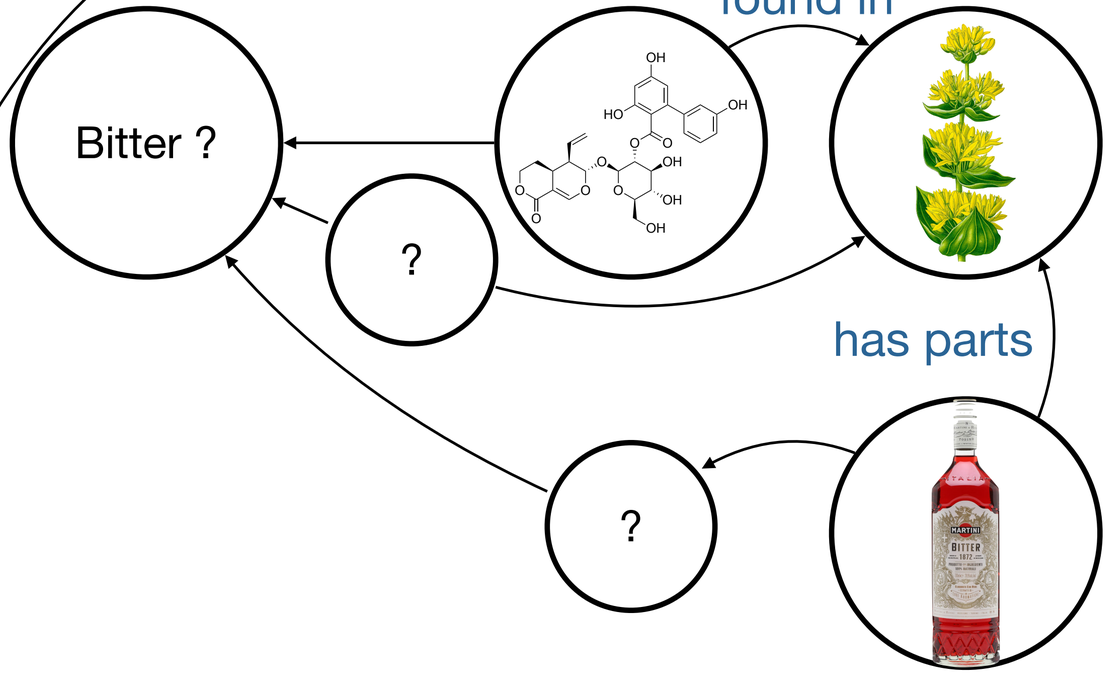


MS-based dereplication

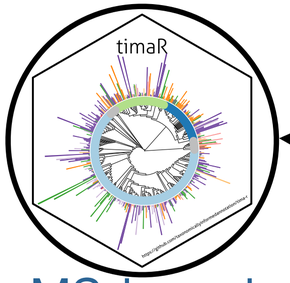


NMR-based dereplication

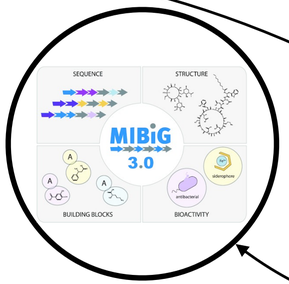




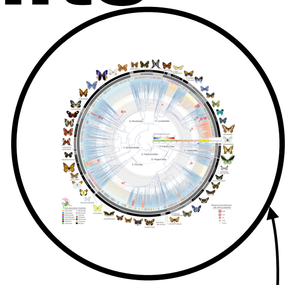
LO(R)D has no limits



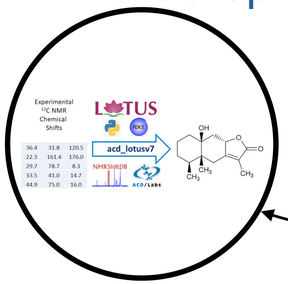
MS-based dereplication



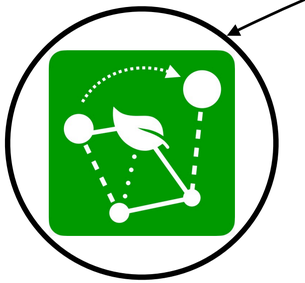
Biosynthetic Gene Clusters



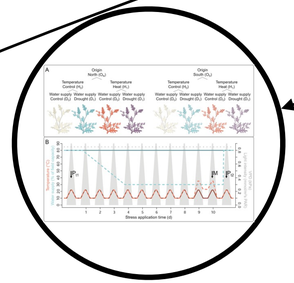
Butterflies



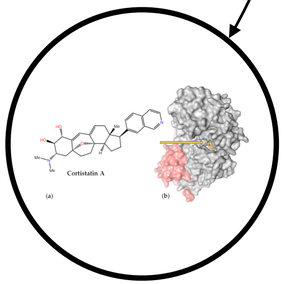
NMR-based dereplication



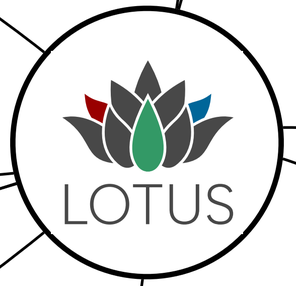
ENPKG

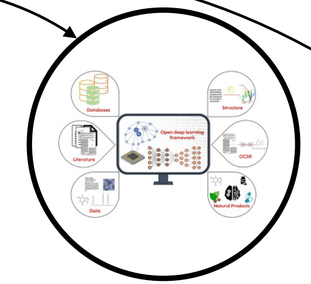


Plant adaptation to heat & drought

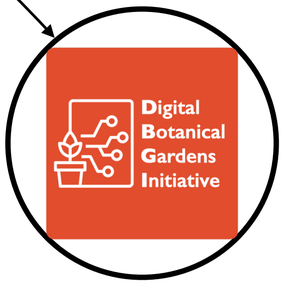


Enzyme inhibition with marine NPs

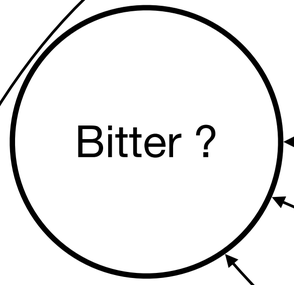


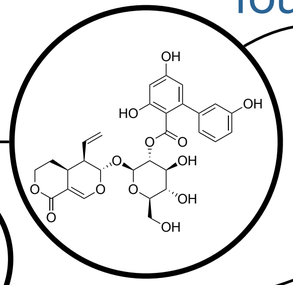


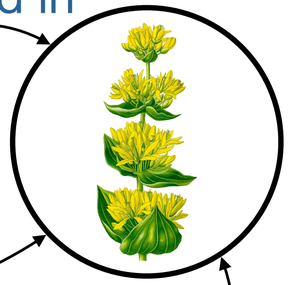
Natural Products & Machine Learning



found in

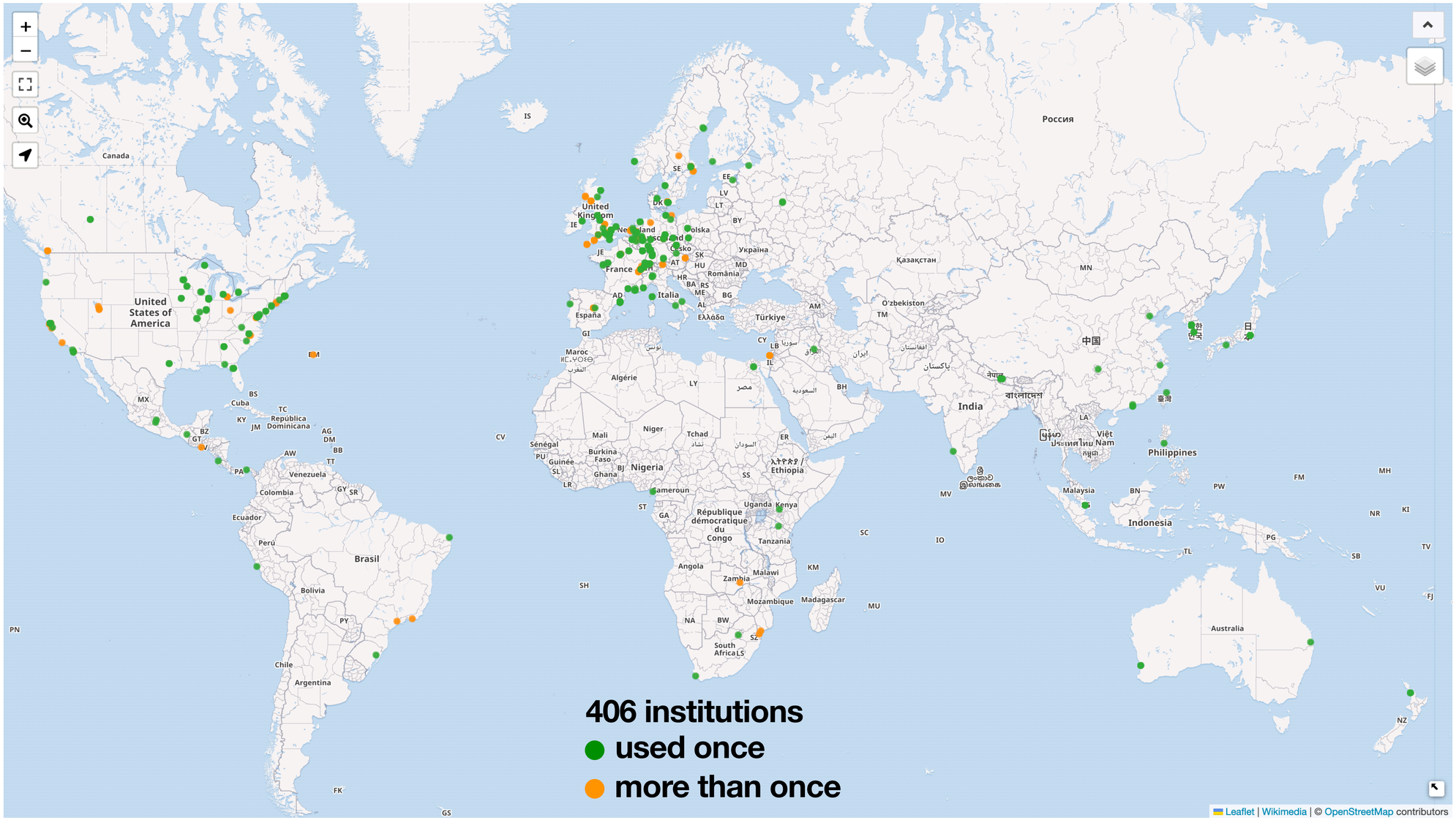






has parts

LO(R)D has no limits



https://w.wiki/A6wV

Acknowledgements

swiss academies of arts and sciences

Swiss National Science Foundation

scnat
akademie der naturwissenschaften
académie des sciences naturelles
swiss academy of sciences
accademia di scienze naturali

WIKIDATA

ETH zürich
D BIOL
Institute of Molecular Systems Biology

Prof. Nicola Zamboni

Contact

rutz@imsb.biol.ethz.ch