ETH zürich

Institute of Molecular Systems Biology

A Distillate of the LOTUS Initiative in Natural Products Research: Sharing Information, Building Knowledge, Extracting Wisdom

71st International Congress and Annual Meeting of the Society for Medicinal Plant and Natural Product Research (GA)

Dublin, 2023-07-03

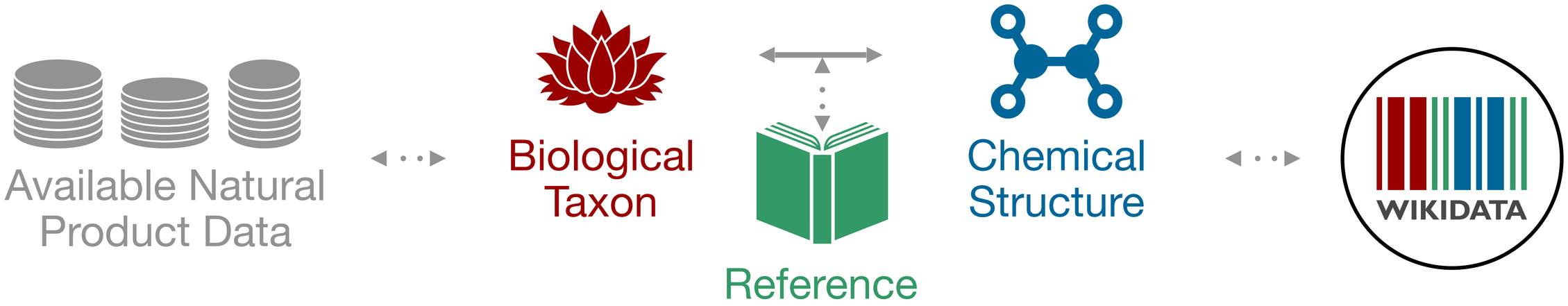
D BIOL

Available Natural Product Data

Unstructured, redundant and noisy data

Unstructured, redundant and noisy data

Cleaned and documented data



Unstructured, redundant and noisy data

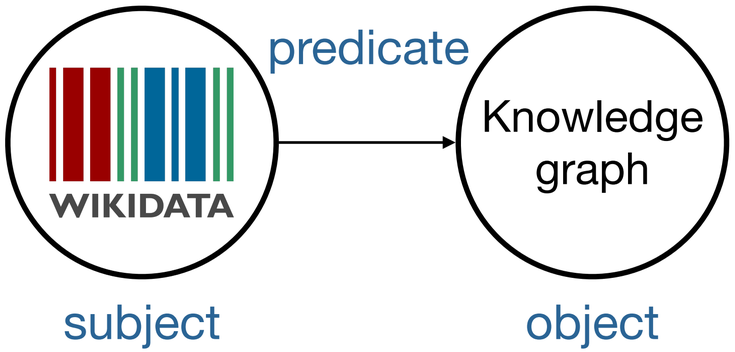
Cleaned and documented data

Collaboratively curated and structured data

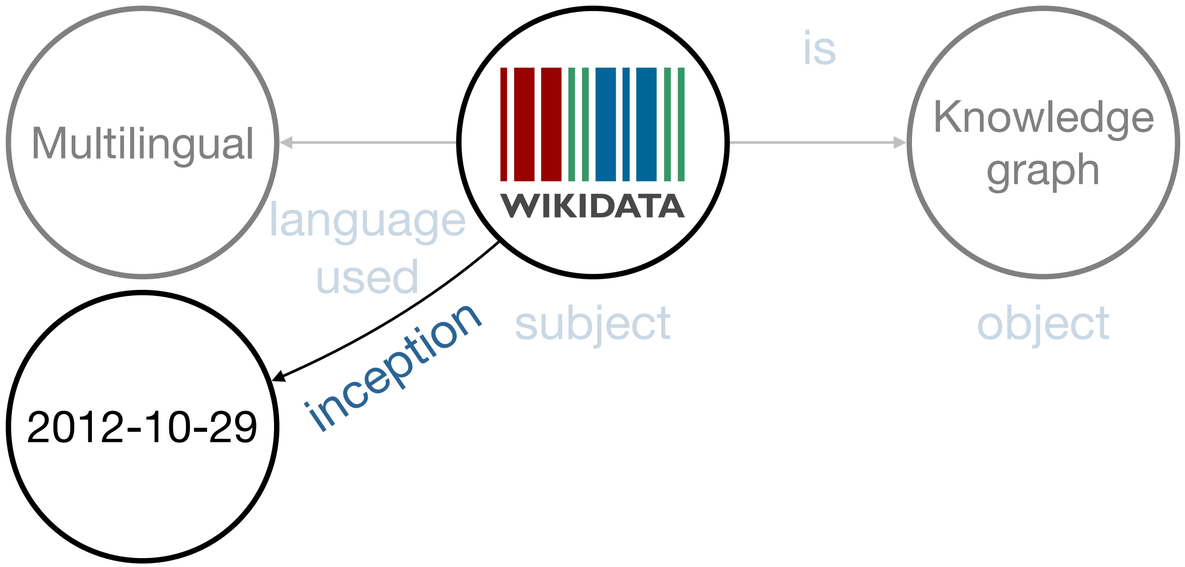
The LOTUS Initiative - **Sharing Information**

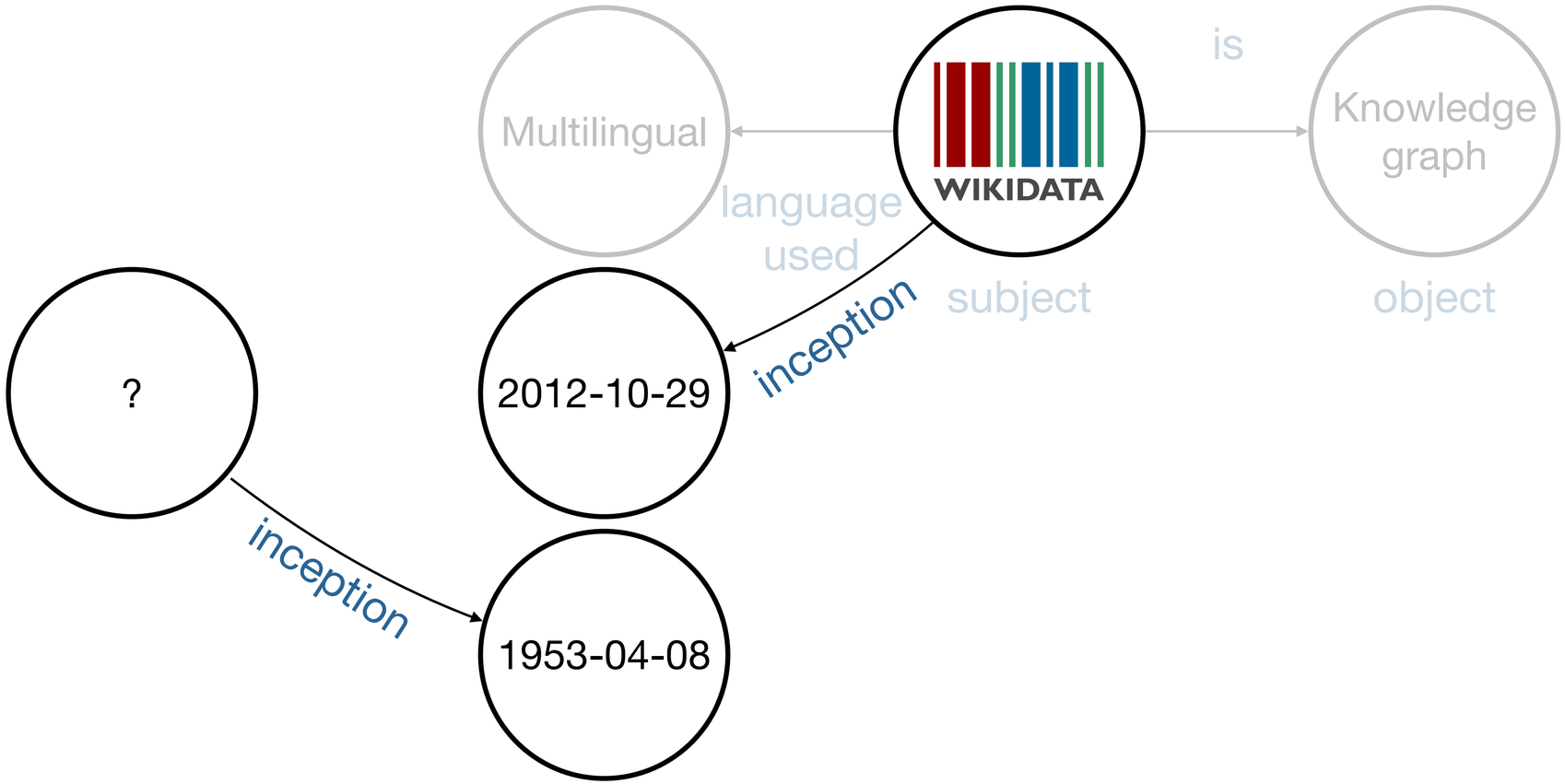
The LOTUS Initiative - **Sharing Information**

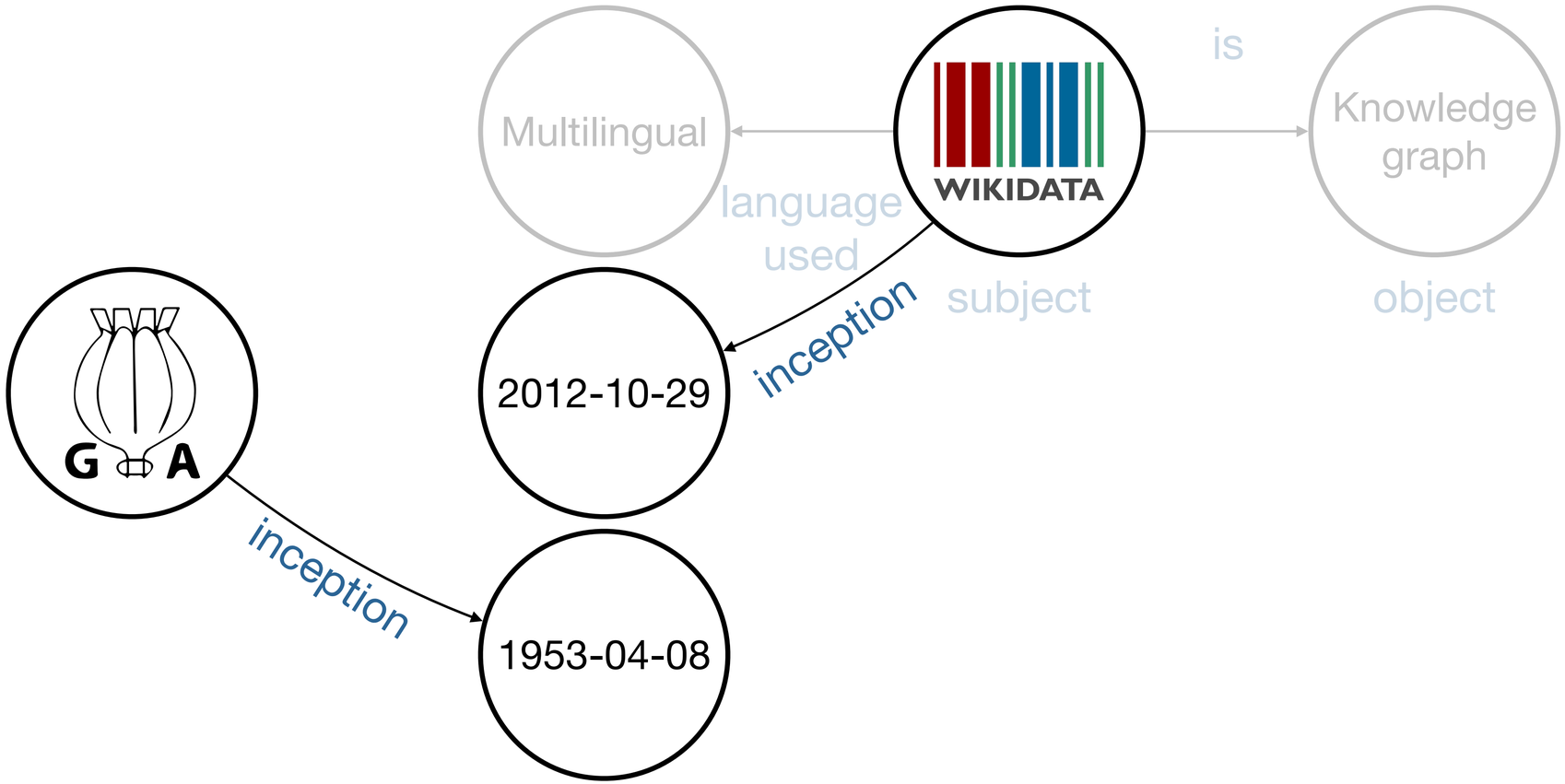
WIKIDATA

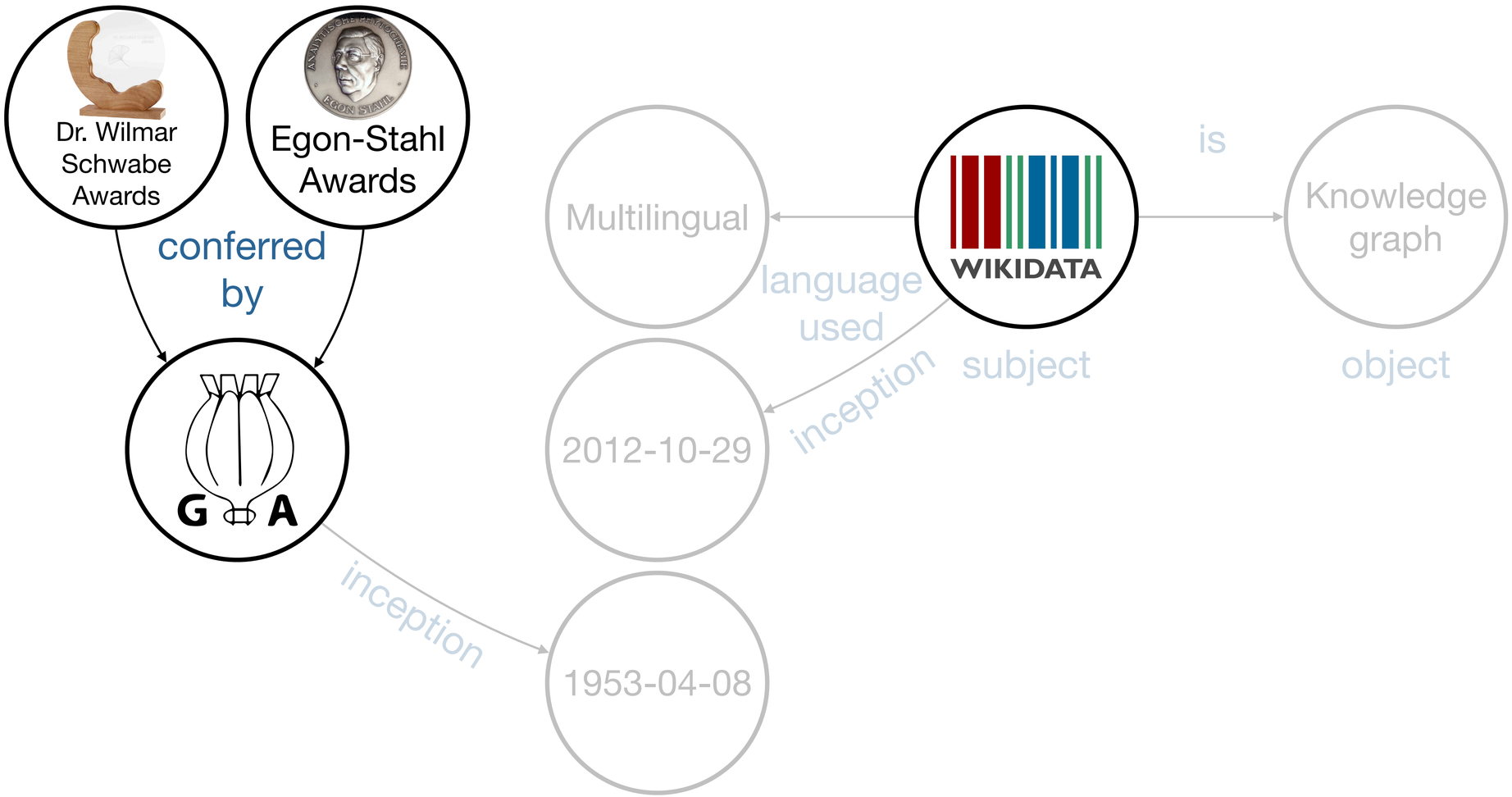

predicate
WIKIDATA
Knowledge graph
subject
object

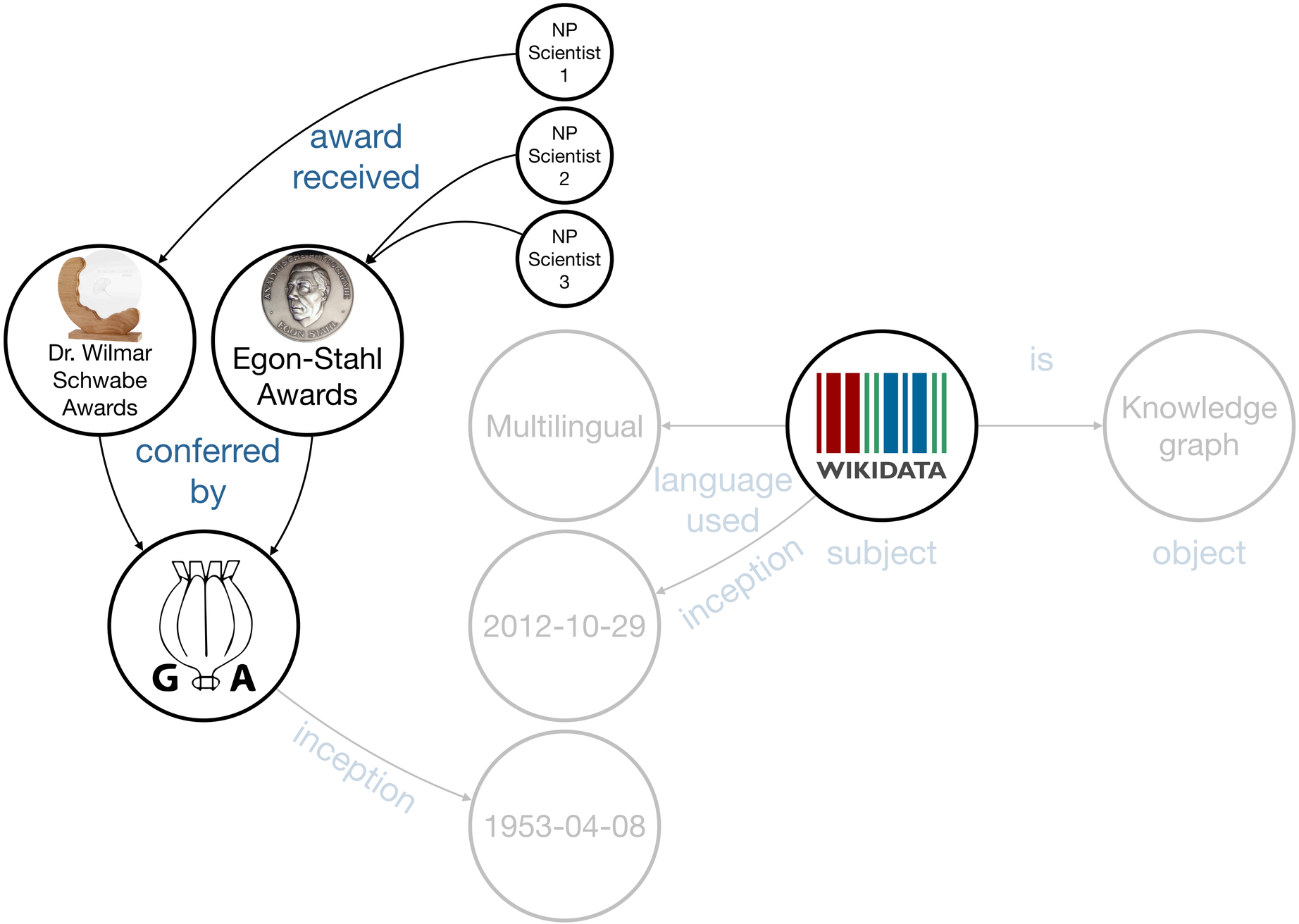
Multilingual
language used
WIKIDATA
subject
is
Knowledge graph
object


Multilingual
WIKIDATA
Knowledge graph
is
language used
subject
object
2012-10-29
inception

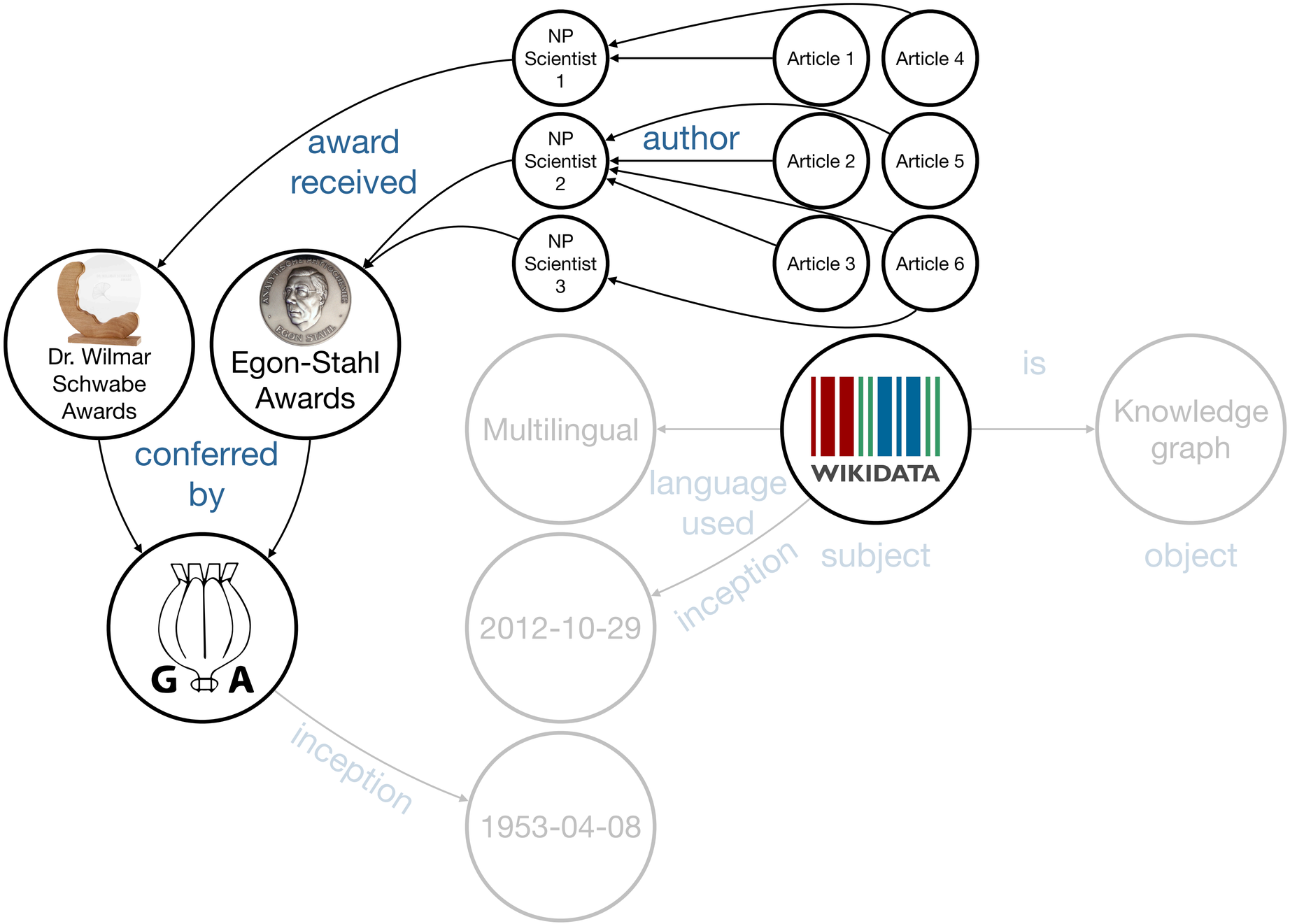

Multilingual
WIKIDATA
is
Knowledge graph
language used
subject
object
?
2012-10-29
inception
inception
1953-04-08


Multilingual
WIKIDATA
is
Knowledge graph
language used
subject
object
2012-10-29
inception
G A
inception
1953-04-08

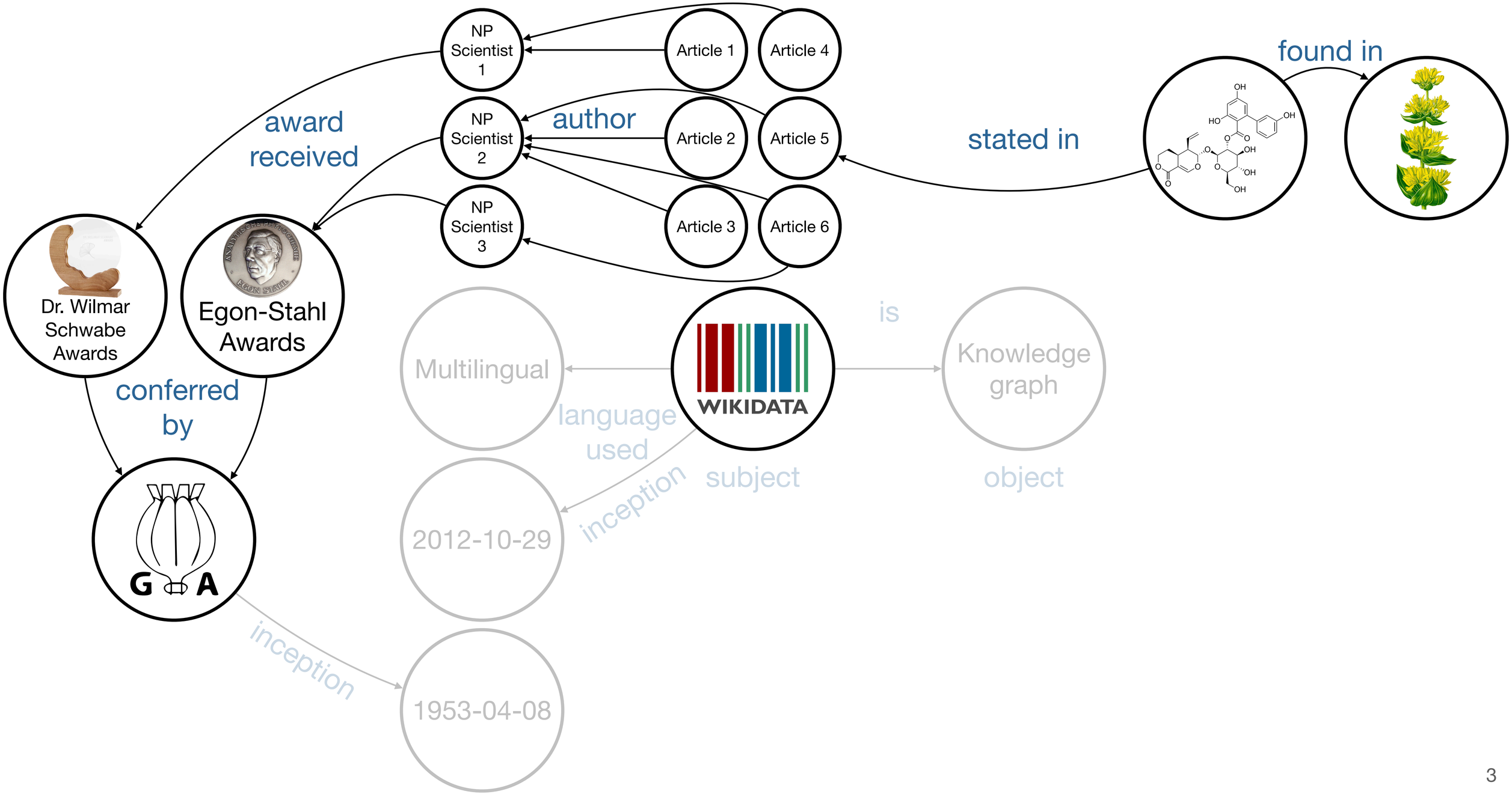

Dr. Wilmar Schwabe Awards
Egon-Stahl Awards
conferred by
G A
inception
1953-04-08
Multilingual
language used
2012-10-29
inception
WIKIDATA
subject
is
Knowledge graph
object


NP Scientist 1
NP Scientist 2
NP Scientist 3
award received
Dr. Wilmar Schwabe Awards
Egon-Stahl Awards
conferred by
G A
inception
1953-04-08
Multilingual
language used
WIKIDATA
is
Knowledge graph
2012-10-29
inception
subject
object

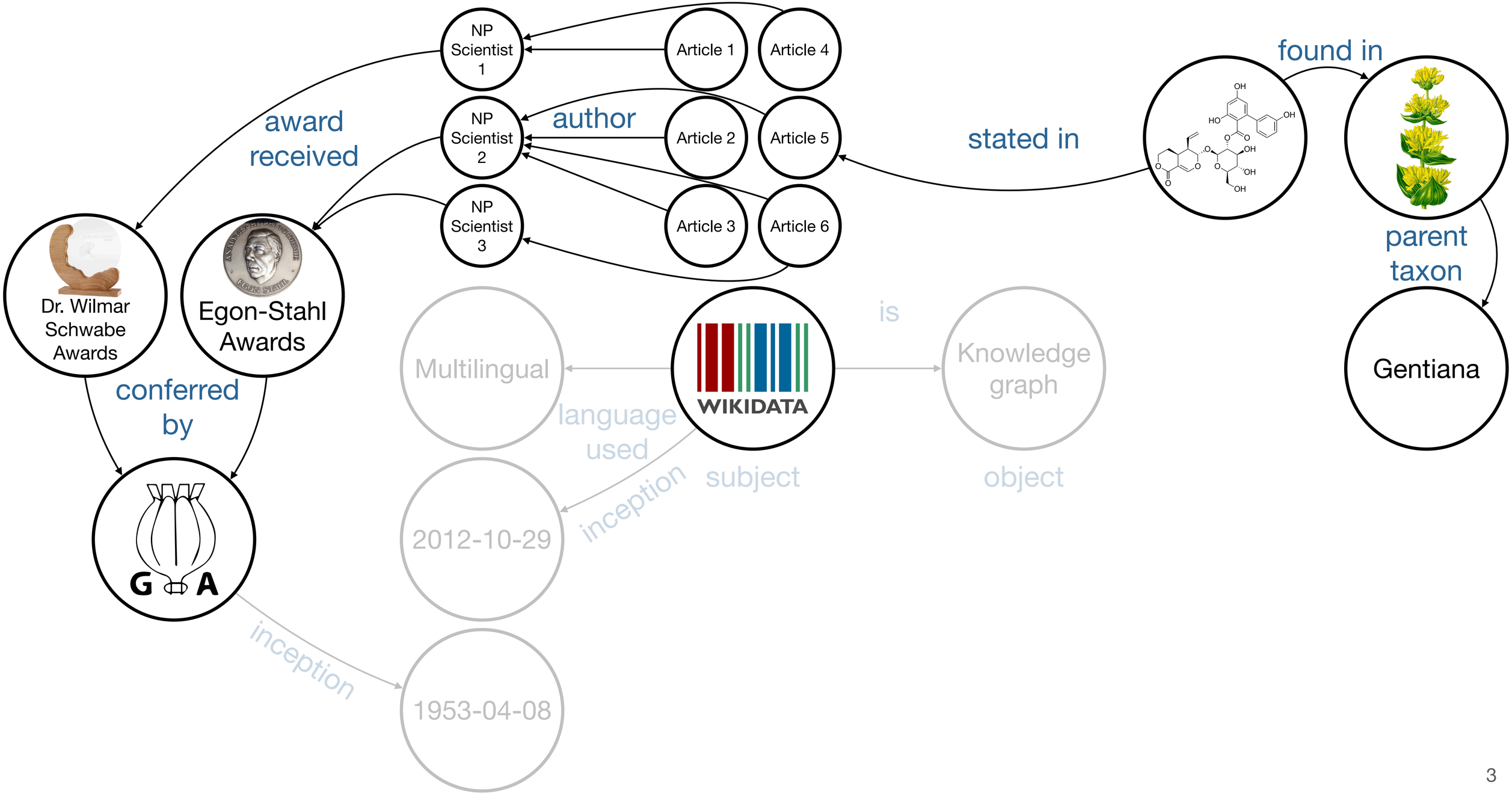
The LOTUS Initiative - **Building Knowledge**

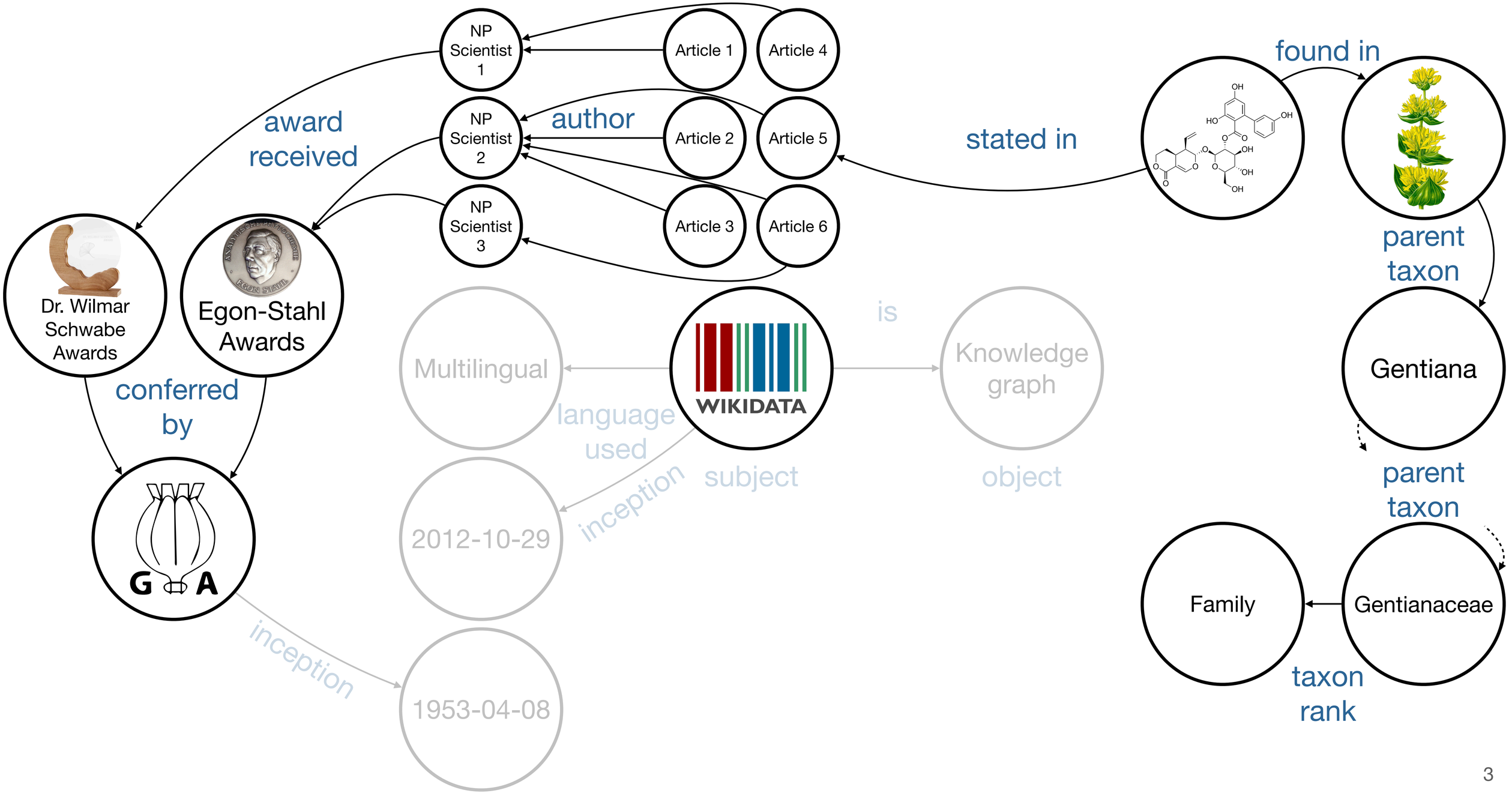


The LOTUS Initiative - **Building Knowledge**

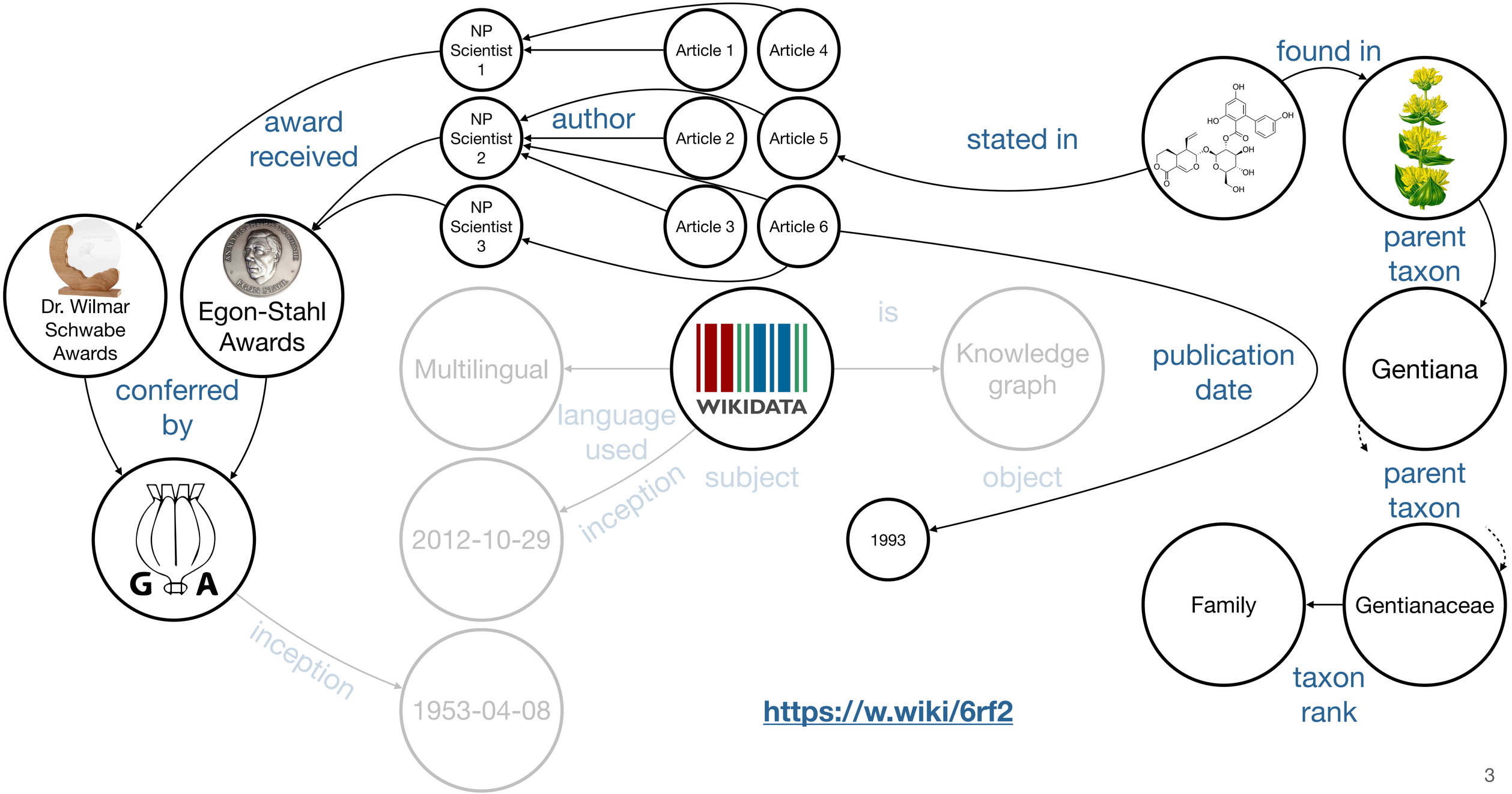


The LOTUS Initiative - **Building Knowledge**

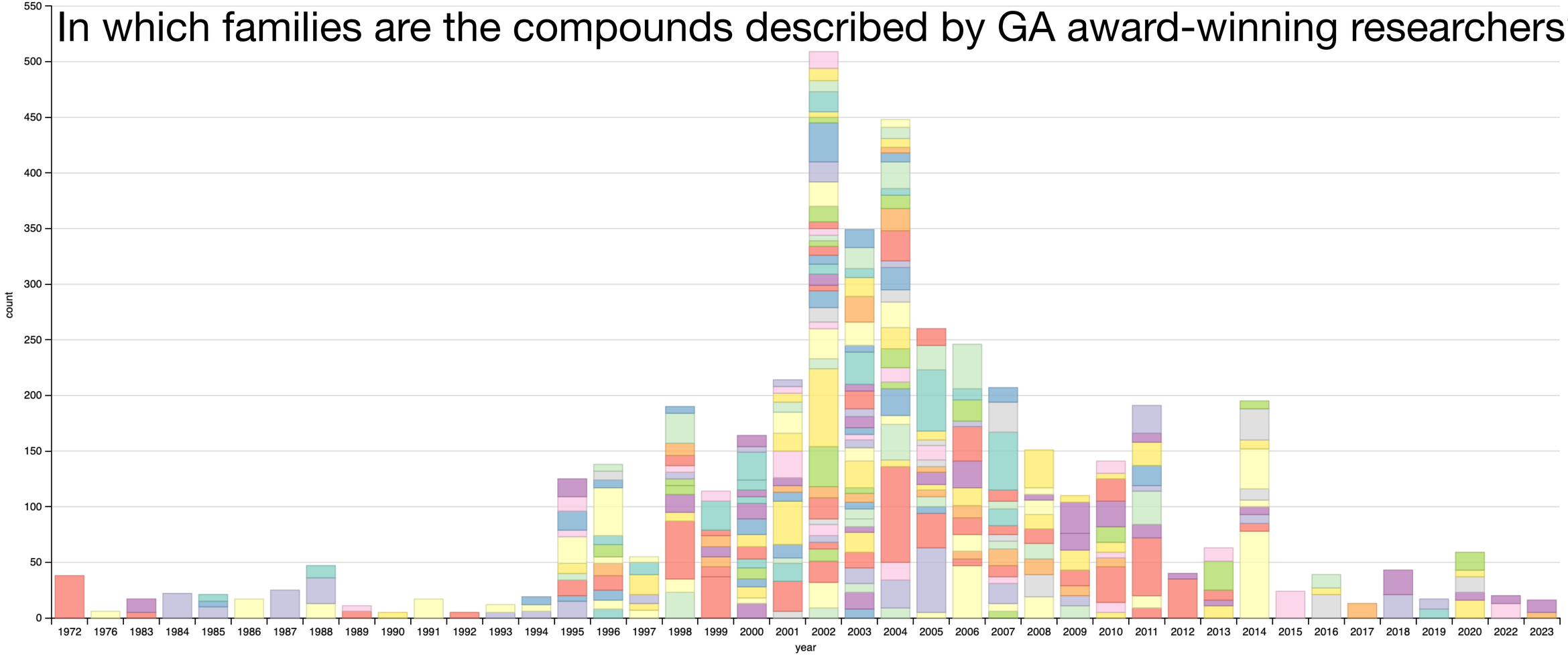



NP Scientist 1
NP Scientist 2
NP Scientist 3
Article 1
Article 2
Article 3
Article 4
Article 5
Article 6
author
award received
stated in
found in
parent taxon
parent taxon
Dr. Wilmar Schwabe Awards
Egon-Stahl Awards
conferred by
WIKIDATA
is
Multilingual
language used
Knowledge graph
subject
object
inception
2012-10-29
inception
1953-04-08
Gentiana
Family
Gentianaceae
taxon rank

The LOTUS Initiative - **Building Knowledge**



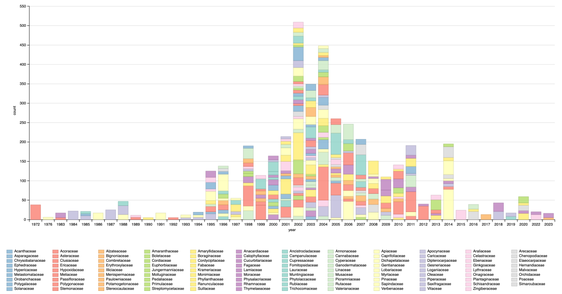
In which families are the compounds described by GA award-winning researchers?



Acanthaceae, Asparagaceae, Chrysobalanaceae, Ephedraceae, Hypericaceae, Melastomataceae, Orobanchaceae, Polygalaceae, Solanaceae, Acoraceae, Asteraceae, Clusiaceae, Ericaceae, Hypoxidaceae, Meliaceae, Passifloraceae, Polygonaceae, Stemonaceae, Alzateaceae, Bignoniaceae, Combretaceae, Erythroxylaceae, Illiciaceae, Menispermaceae, Paulowniaceae, Potamogetonaceae, Stereocaulaceae, Amaranthaceae, Boletaceae, Cordiaceae, Euphorbiaceae, Jungermanniaceae, Molluginaceae, Pedaliaceae, Primulaceae, Streptomycetaceae, Amaryllidaceae, Boraginaceae, Cordycipitaceae, Fabaceae, Krameriaceae, Monimiaceae, Phyllanthaceae, Ranunculaceae, Suillaceae, Anacardiaceae, Calophyllaceae, Cucurbitariaceae, Fagaceae, Lamiaceae, Moraceae, Physalacriaceae, Rhamnaceae, Thymelaeaceae, Ancistrocladaceae, Campanulaceae, Cupressaceae, Fomitopsidaceae, Lauraceae, Muntingiaceae, Phytolaccaceae, Rubiaceae, Trichocomaceae, Annonaceae, Cannabaceae, Cyperaceae, Ganodermataceae, Linaceae, Musaceae, Picramniaceae, Rutaceae, Valerianaceae, Apiaceae, Caprifoliaceae, Dichapetalaceae, Gentianaceae, Lobariaceae, Myrtaceae, Pinaceae, Sapindaceae, Verbenaceae, Apocynaceae, Caricaceae, Dipterocarpaceae, Gesneriaceae, Loganiaceae, Oleaceae, Piperaceae, Saxifragaceae, Vitaceae, Araliaceae, Celastraceae, Ebenaceae, Ginkgoaceae, Lythraceae, Onagraceae, Plantaginaceae, Schisandraceae, Zingiberaceae, Arecaceae, Chenopodiaceae, Elaeocarpaceae, Hernandiaceae, Malvaceae, Orchidaceae, Poaceae, Simaroubaceae

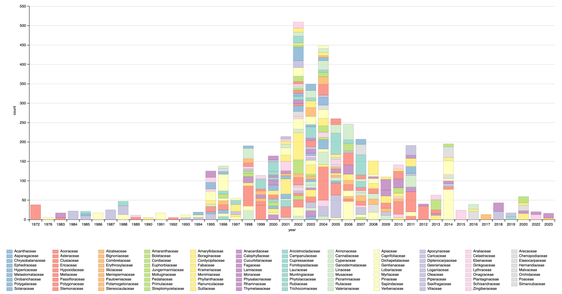
https://w.wiki/6rf2

The LOTUS Initiative - a **Collaborative Platform**

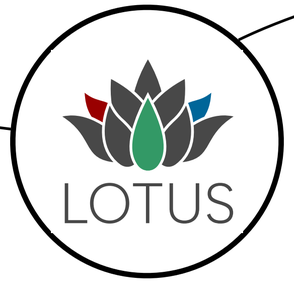


https://w.wiki/6rf2

The LOTUS Initiative - a **Collaborative Platform**



https://w.wiki/6rf2



WIKIDATA

Item Discussion Read View history

Main page
Community portal
Project chat
Create a new Item
Recent changes
Random Item
Query Service
Nearby
Help
Donate

Lexicographical data
Create a new Lexeme
Recent changes
Random Lexeme

Tools
What links here
Related changes
Special pages
Permanent link
Page information
Concept URI
Cite this page

hyperelodione D (Q116482353)

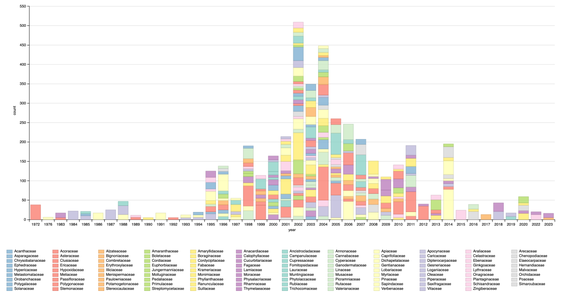
chemical compound

Language	Label	Description
English	hyperelodione D	chemical compound

Statements

instance of	chemical compound 0 references
mass	548.797 dalton 1 reference based on heuristic: inferred from SMILES
chemical formula	$C_{36}H_{52}O_4$ 1 reference based on heuristic: inferred from SMILES

The LOTUS Initiative - a **Collaborative Platform**



https://w.wiki/6rf2

WIKIDATA

Item Discussion Read View history

Main page
Community portal
Project chat
Create a new Item
Recent changes
Random Item
Query Service
Nearby
Help
Donate

Lexicographical data
Create a new Lexeme
Recent changes
Random Lexeme

Tools
What links here
Related changes
Special pages
Permanent link
Page information
Concept URI
Cite this page

hyperelodione D (Q116482353)

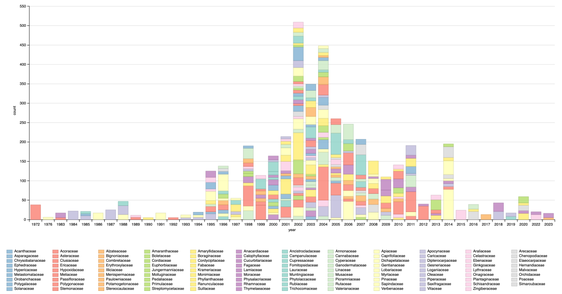
chemical compound

Language	Label	Description
English	hyperelodione D	chemical compound

Statements

instance of	chemical compound 0 references
mass	548.797 dalton 1 reference based on heuristic: inferred from SMILES
chemical formula	$C_{36}H_{52}O_4$ 1 reference based on heuristic: inferred from SMILES
canonical SMILES	C/C(C)=C\CC/C(C)=C/CC1=C2C3(O[H])C(C4(C/C=C(C)/CC/C=C(C)/C)C1=O)(O[H])C(C(C)=CC4)([H])CC3([H])C(C)(C)O2 reason for deprecated rank: structural reassignment 1 reference stated in: Cytotoxic polyprenylated phloroglucinol derivatives from Hypericum elodeoides Choisy modulating the transactivation of RXRα
	C/C(C)=C/CC1=C2C3(O[H])C(C4(C/C=C(C)/CC/C=C(C)/C)C1=O)(O[H])C(C(C)=CC4)([H])CC3([H])C(CC/C=C(C)/C)(C)O2 reason for preferred rank: structural reassignment 1 reference stated in: Bioinspired Total Synthesis of Erectones A and B, and the Revised Structure of Hyperelodione D

The LOTUS Initiative - a **Collaborative Platform**



https://w.wiki/6rf2

Universität Basel

LOTUS

WIKIDATA

Item | Discussion | Read | View history

hyperelodione D (Q116482353)

chemical compound

Main page
Community portal
Project chat
Create a new Item
Recent changes
Random Item
Query Service
Nearby
Help
Donate

Lexicographical data
Create a new Lexeme
Recent changes
Random Lexeme

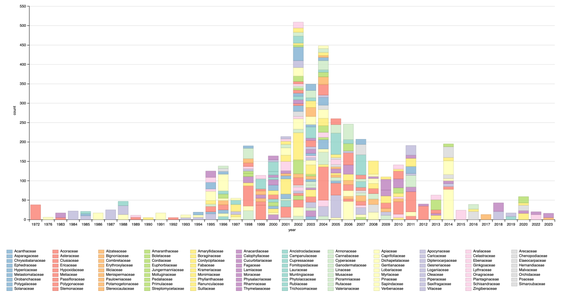
Tools
What links here
Related changes
Special pages
Permanent link
Page information
Concept URI
Cite this page

Language	Label	Description
English	hyperelodione D	chemical compound

Statements

instance of	chemical compound 0 references
mass	548.797 dalton 1 reference based on heuristic: inferred from SMILES
chemical formula	$C_{36}H_{52}O_4$ 1 reference based on heuristic: inferred from SMILES
canonical SMILES	C/C(C)=C\CC/C(C)=C/CC1=C2C3(O[H])C(C4(C/C=C(C)/CC/C=C(C)/C)C1=O)(O[H])C(C(C)=CC4)([H])CC3([H])C(C)(C)O2 reason for deprecated rank: structural reassignment 1 reference stated in: Cytotoxic polyprenylated phloroglucinol derivatives from Hypericum elodeoides Choisy modulating the transactivation of RXRα
	C/C(C)=C/CC1=C2C3(O[H])C(C4(C/C=C(C)/CC/C=C(C)/C)C1=O)(O[H])C(C(C)=CC4)([H])CC3([H])C(CC/C=C(C)/C)(C)O2 reason for preferred rank: structural reassignment 1 reference stated in: Bioinspired Total Synthesis of Erectones A and B, and the Revised Structure of Hyperelodione D

The LOTUS Initiative - a **Collaborative Platform**



https://w.wiki/6rf2

Universität Basel

University of Manitoba

LOTUS

WIKIDATA

Item Discussion Read View history

hyperelodione D (Q116482353)

chemical compound

Main page
Community portal
Project chat
Create a new Item
Recent changes
Random Item
Query Service
Nearby
Help
Donate

Lexicographical data
Create a new Lexeme
Recent changes
Random Lexeme

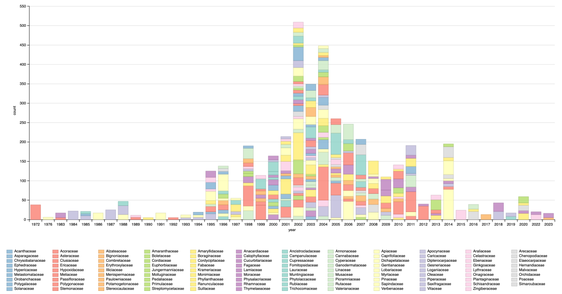
Tools
What links here
Related changes
Special pages
Permanent link
Page information
Concept URI
Cite this page

Language	Label	Description
English	hyperelodione D	chemical compound

Statements

instance of	chemical compound 0 references
mass	548.797 dalton 1 reference based on heuristic: inferred from SMILES
chemical formula	$C_{36}H_{52}O_4$ 1 reference based on heuristic: inferred from SMILES
canonical SMILES	C/C(C)=C\CC/C(C)=C/CC1=C2C3(O[H])C(C4(C/C=C(C)/CC/C=C(C)/C)C1=O)(O[H])C(C(C)=CC4)([H])CC3([H])C(C)(C)O2 reason for deprecated rank: structural reassignment 1 reference stated in: Cytotoxic polyprenylated phloroglucinol derivatives from Hypericum elodeoides Choisy modulating the transactivation of RXRα
	C/C(C)=C/CC1=C2C3(O[H])C(C4(C/C=C(C)/CC/C=C(C)/C)C1=O)(O[H])C(C(C)=CC4)([H])CC3([H])C(CC/C=C(C)/C)(C)O2 reason for preferred rank: structural reassignment 1 reference stated in: Bioinspired Total Synthesis of Erectones A and B, and the Revised Structure of Hyperelodione D

The LOTUS Initiative - a **Collaborative Platform**



https://w.wiki/6rf2

Universität Basel

University of Manitoba

université PARIS-SACLAY

LOTUS

WIKIDATA

Item | Discussion | Read | View history

hyperelodione D (Q116482353)

chemical compound

Main page
Community portal
Project chat
Create a new Item
Recent changes
Random Item
Query Service
Nearby
Help
Donate

Lexicographical data
Create a new Lexeme
Recent changes
Random Lexeme

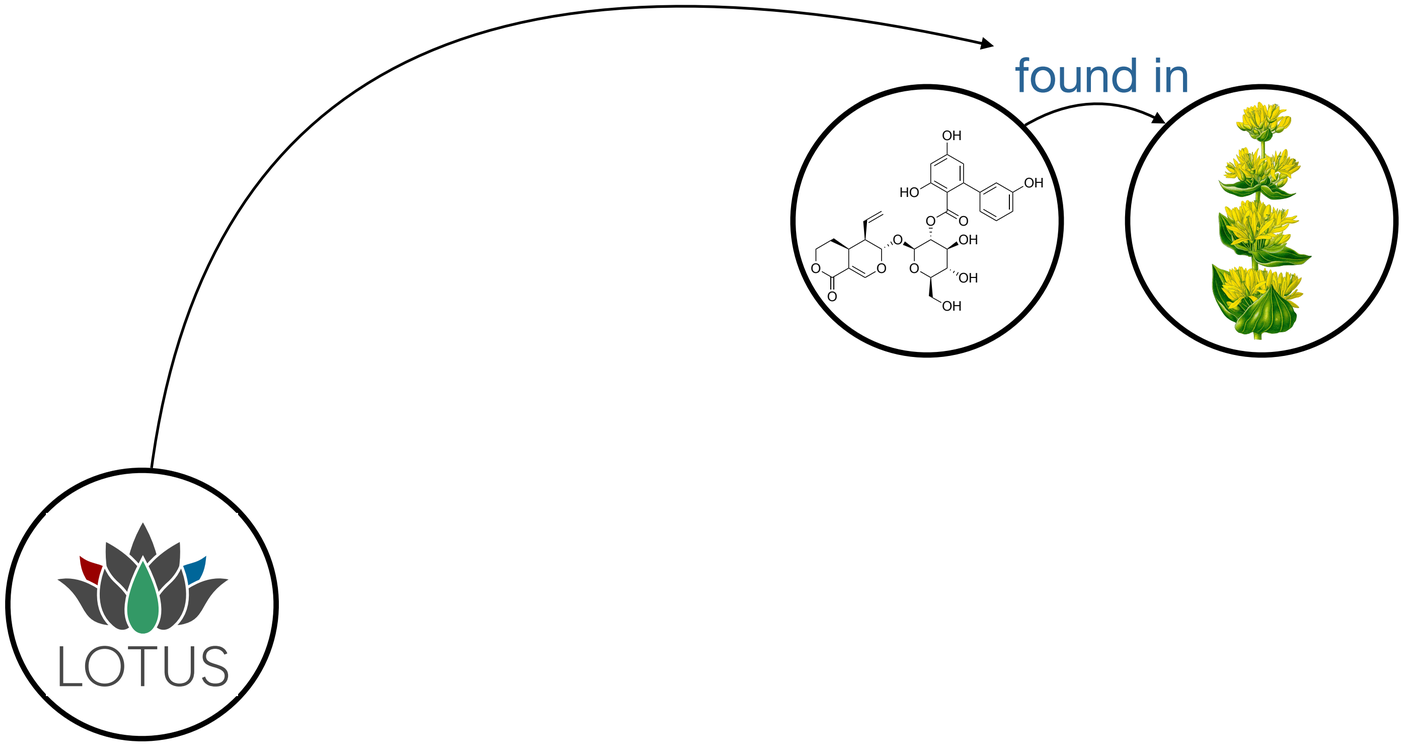
Tools
What links here
Related changes
Special pages
Permanent link
Page information
Concept URI
Cite this page

Language	Label	Description
English	hyperelodione D	chemical compound

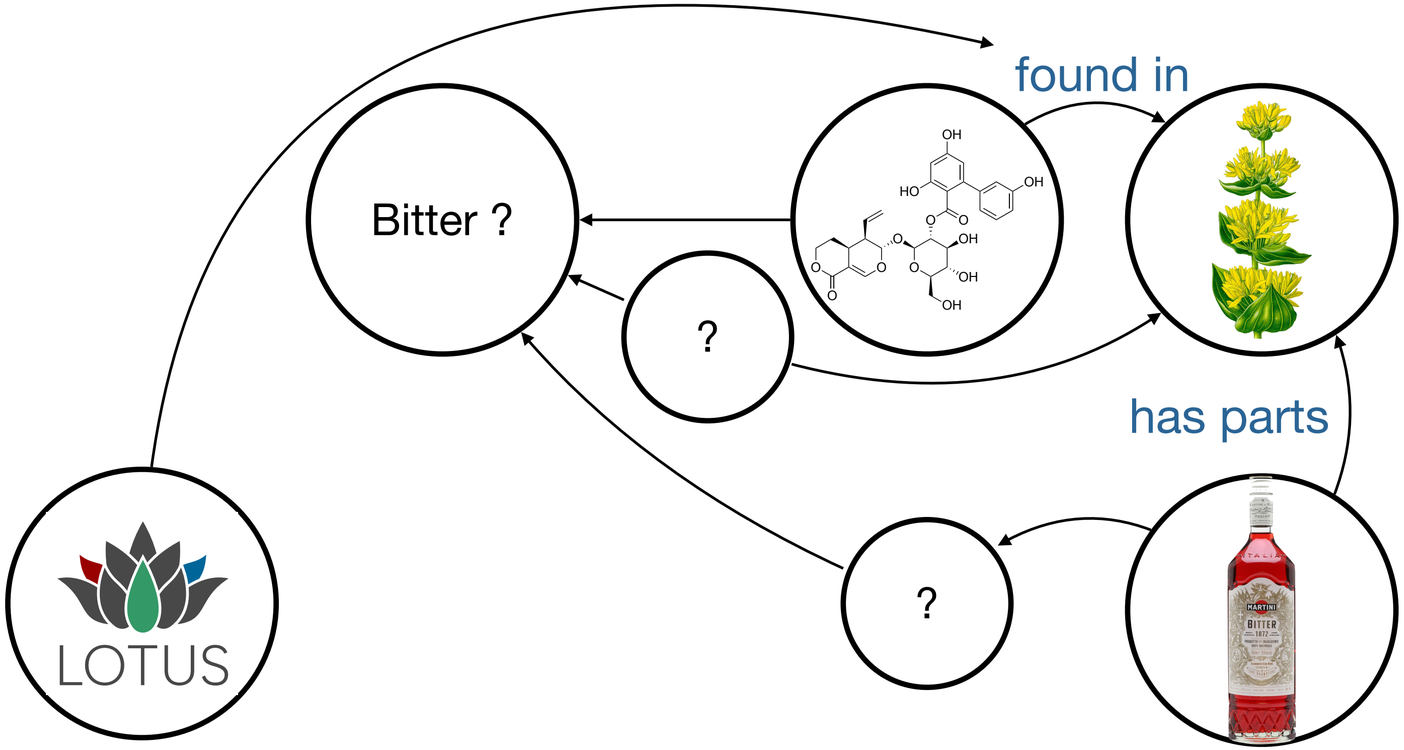
Statements

instance of	chemical compound 0 references
mass	548.797 dalton 1 reference based on heuristic: inferred from SMILES
chemical formula	$C_{36}H_{52}O_4$ 1 reference based on heuristic: inferred from SMILES
canonical SMILES	C/C(C)=C\CC/C(C)=C/CC1=C2C3(O[H])C(C4(C/C=C(C)/CC/C=C(C)/C)C1=O)(O[H])C(C(C)=CC4)([H])CC3([H])C(C)(C)O2 reason for deprecated rank: structural reassignment 1 reference stated in: Cytotoxic polyprenylated phloroglucinol derivatives from Hypericum elodeoides Choisy modulating the transactivation of RXRα
	C/C(C)=C/CC1=C2C3(O[H])C(C4(C/C=C(C)/CC/C=C(C)/C)C1=O)(O[H])C(C(C)=CC4)([H])CC3([H])C(CC/C=C(C)/C)(C)O2 reason for preferred rank: structural reassignment 1 reference stated in: Bioinspired Total Synthesis of Erectones A and B, and the Revised Structure of Hyperelodione D

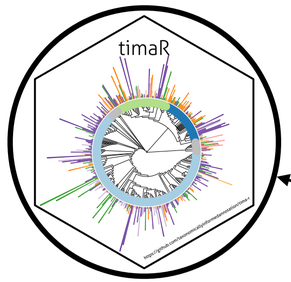
The LOTUS Initiative - **Extracting Wisdom**



The LOTUS Initiative - **Extracting Wisdom**



The LOTUS Initiative - **Extracting Wisdom**

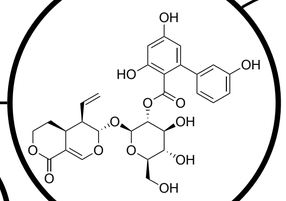


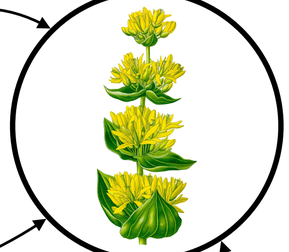
MS-based dereplication

found in

Bitter ?

?

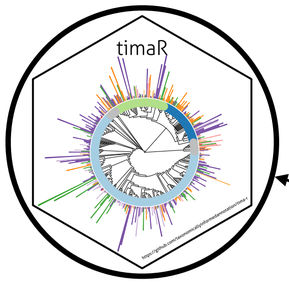




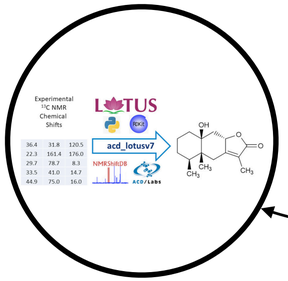
has parts

?

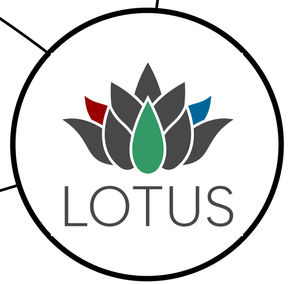
The LOTUS Initiative - **Extracting Wisdom**



MS-based dereplication

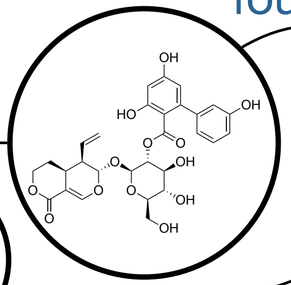


NMR-based dereplication

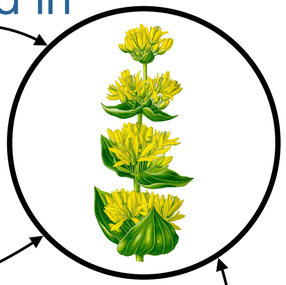


found in

Bitter ?



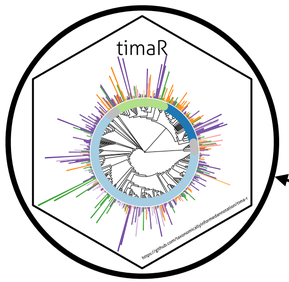
?



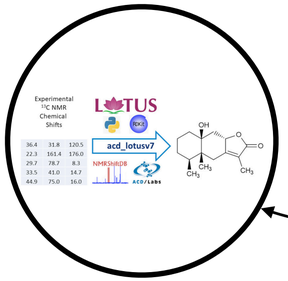
has parts

?

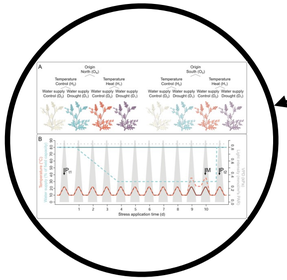
The LOTUS Initiative - **Extracting Wisdom**



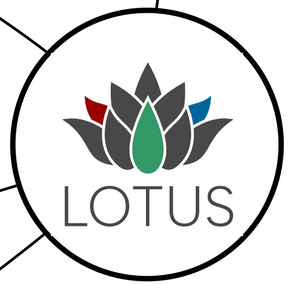
MS-based dereplication



NMR-based dereplication



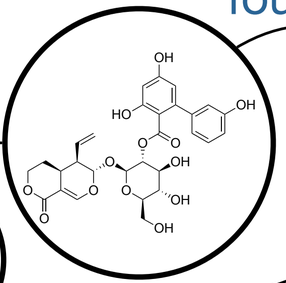
Plant adaptation to heat & drought

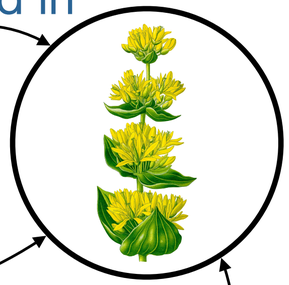


found in

Bitter ?

?

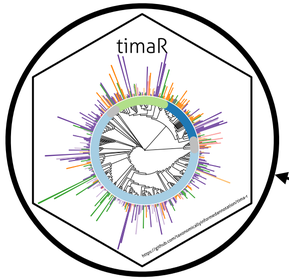




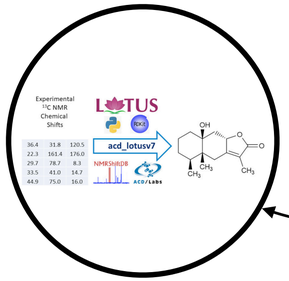
has parts

?

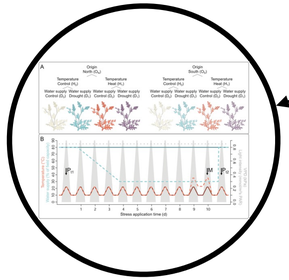
The LOTUS Initiative - **Extracting Wisdom**



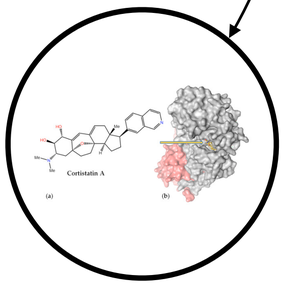
MS-based dereplication



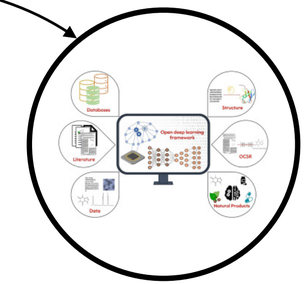
NMR-based dereplication



Plant adaptation to heat & drought



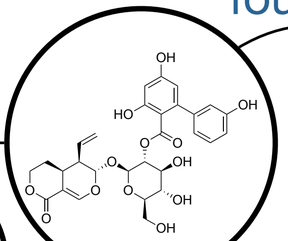
Enzyme inhibition with marine NPs



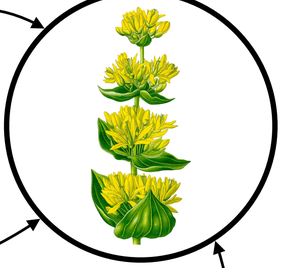
Natural Products & Machine Learning

found in

Bitter ?

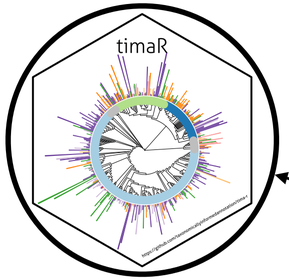


?

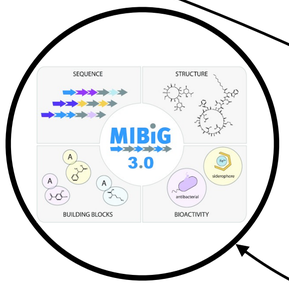


has parts

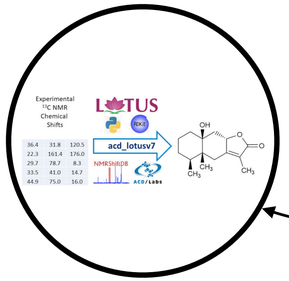
?



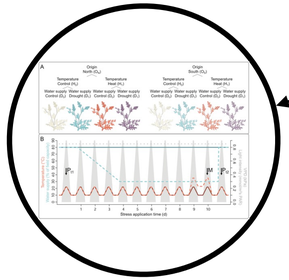
MS-based dereplication



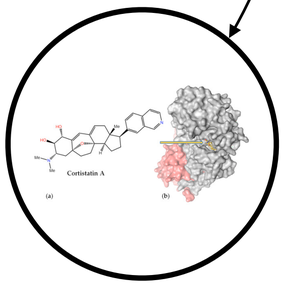
Biosynthetic Gene Clusters



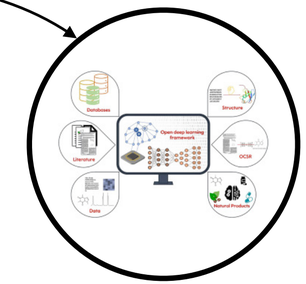
NMR-based dereplication



Plant adaptation to heat & drought



Enzyme inhibition with marine NPs

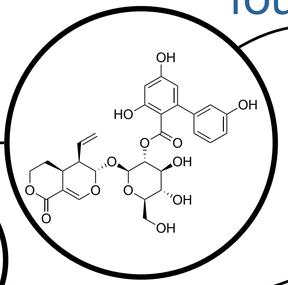


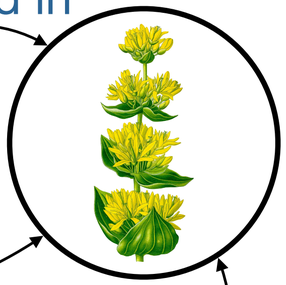
Natural Products & Machine Learning

found in

Bitter ?

?

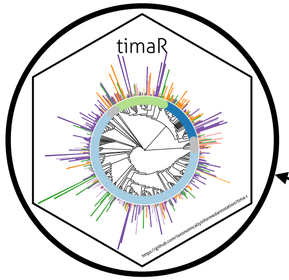




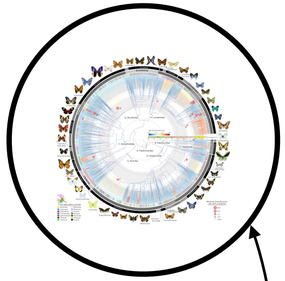
has parts

?

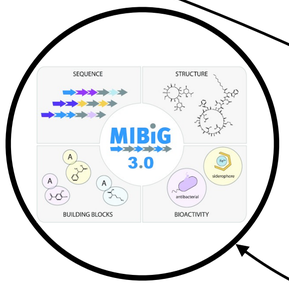
The LOTUS Initiative - **Extracting Wisdom**



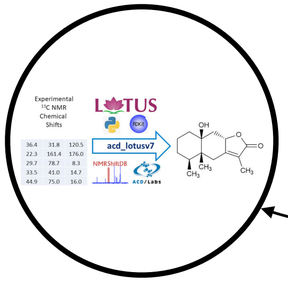
MS-based dereplication



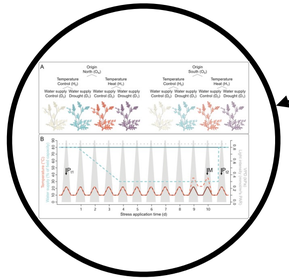
Butterflies phylogeny



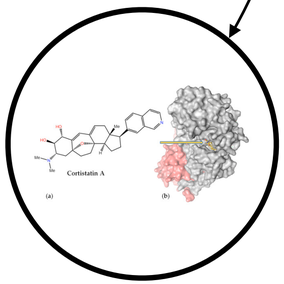
Biosynthetic Gene Clusters



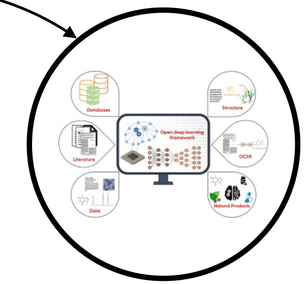
NMR-based dereplication



Plant adaptation to heat & drought



Enzyme inhibition with marine NPs

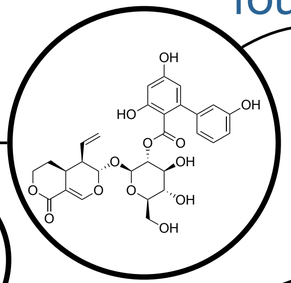


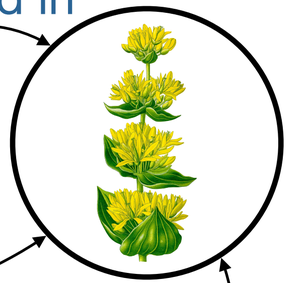
Natural Products & Machine Learning

found in

Bitter ?

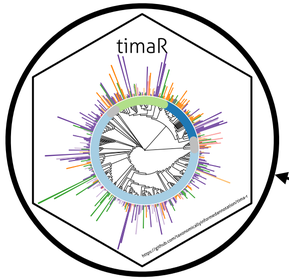
?



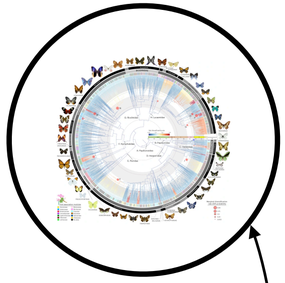


has parts

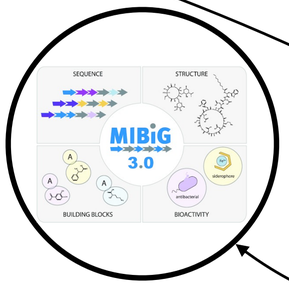
?



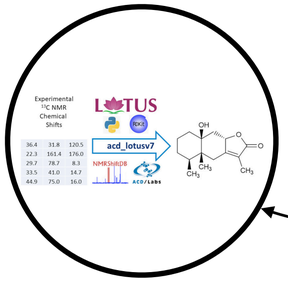
MS-based dereplication



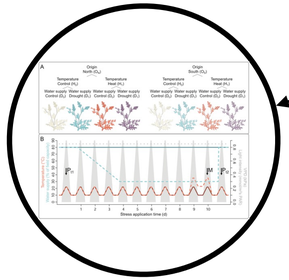
Butterflies phylogeny



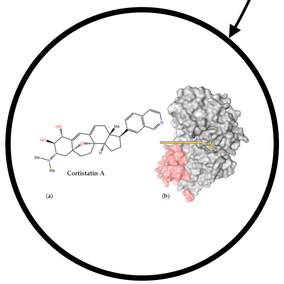
Biosynthetic Gene Clusters



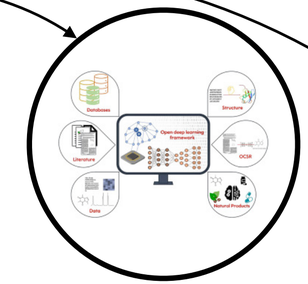
NMR-based dereplication



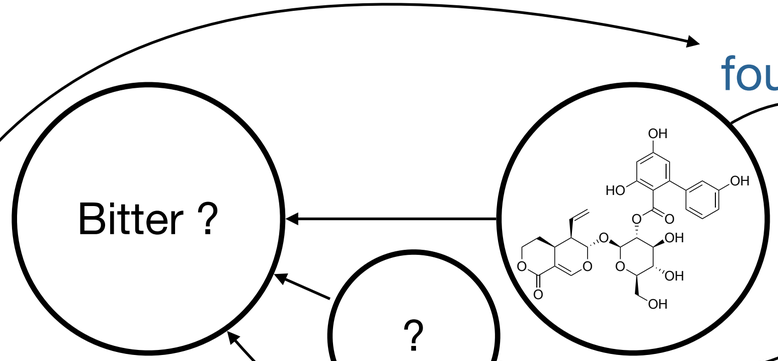
Plant adaptation to heat & drought

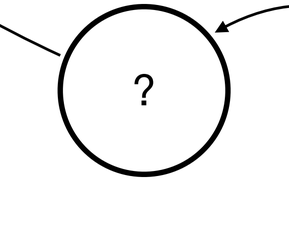


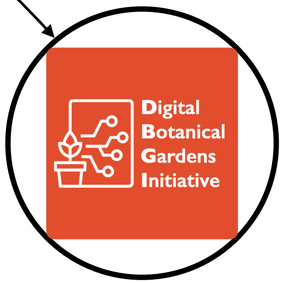
Enzyme inhibition with marine NPs



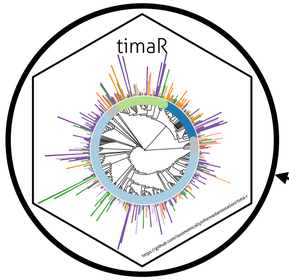
Natural Products & Machine Learning



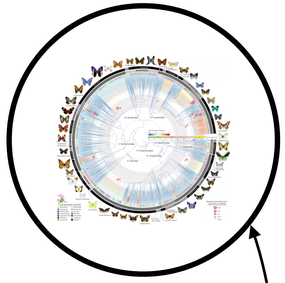




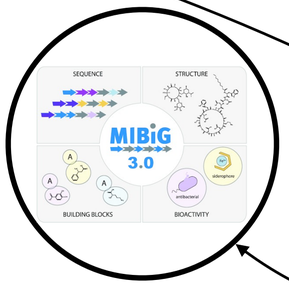
The LOTUS Initiative - **Extracting Wisdom**



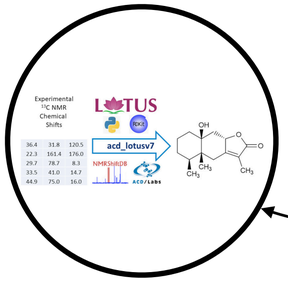
MS-based dereplication



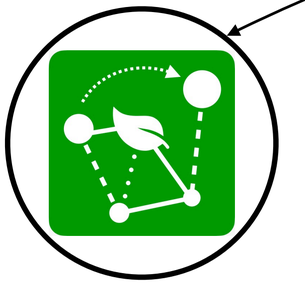
Butterflies phylogeny



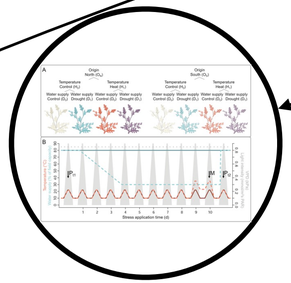
Biosynthetic Gene Clusters



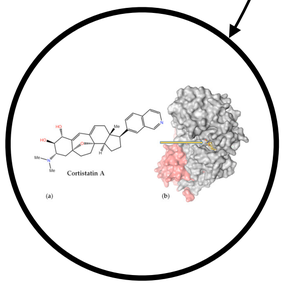
NMR-based dereplication



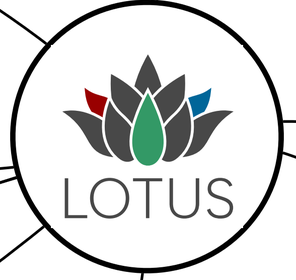
Out today!

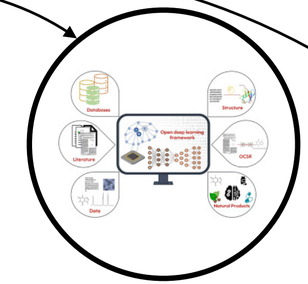


Plant adaptation to heat & drought



Enzyme inhibition with marine NPs



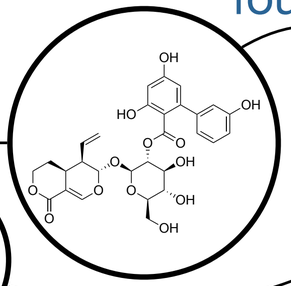


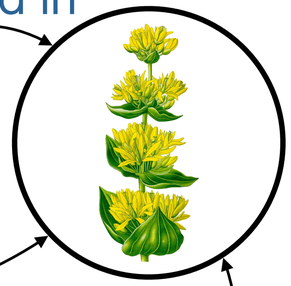
Natural Products & Machine Learning

found in

Bitter ?

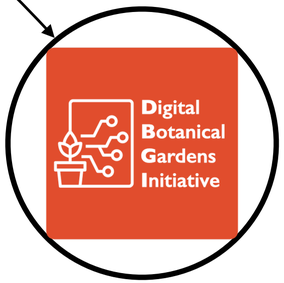
?



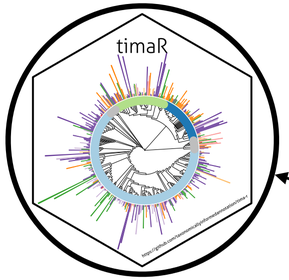


has parts

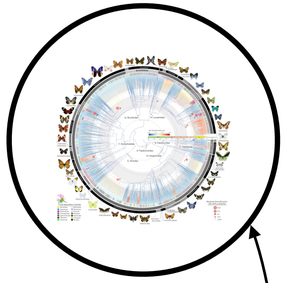
?



The LOTUS Initiative - **Extracting Wisdom**



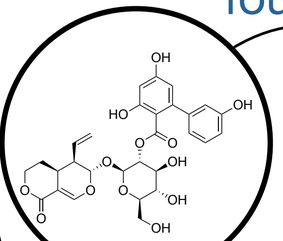
MS-based dereplication



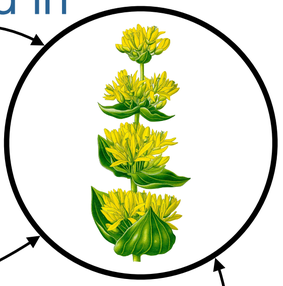
Butterflies phylogeny

found in

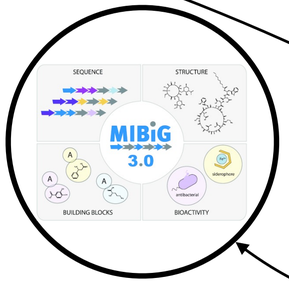
Bitter ?



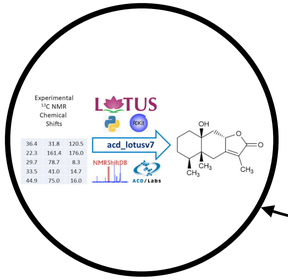
?



has parts



Biosynthetic Gene Clusters

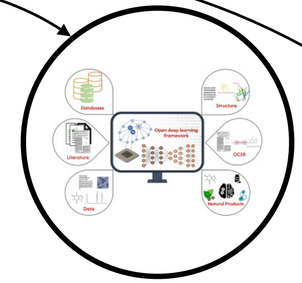


NMR-based dereplication

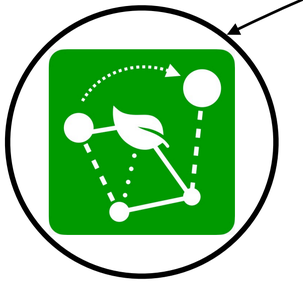
LOTUS

Next ?

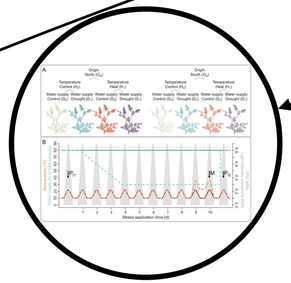
?



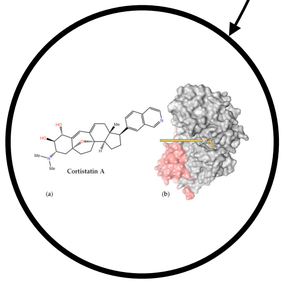
Natural Products & Machine Learning



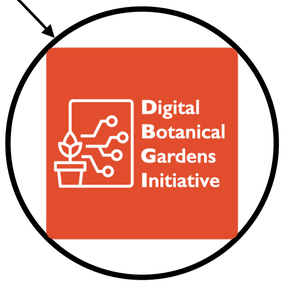
Out today!



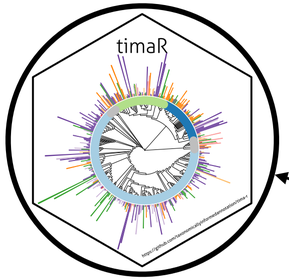
Plant adaptation to heat & drought



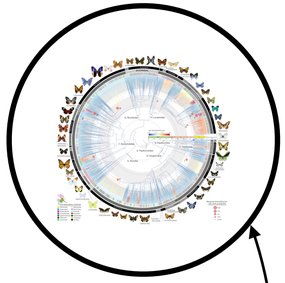
Enzyme inhibition with marine NPs



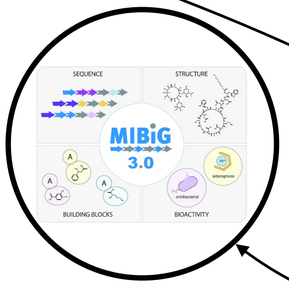
The LOTUS Initiative - **Extracting Wisdom**



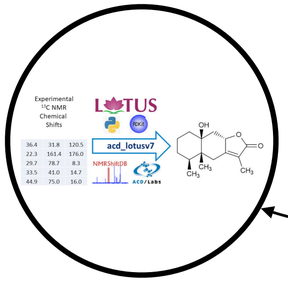
MS-based dereplication



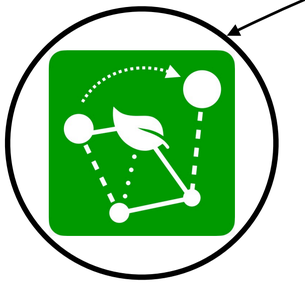
Butterflies phylogeny



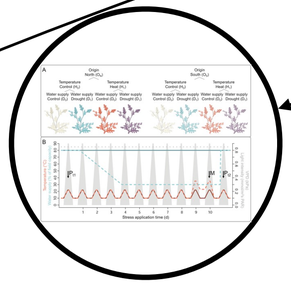
Biosynthetic Gene Clusters



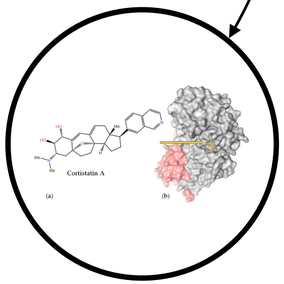
NMR-based dereplication



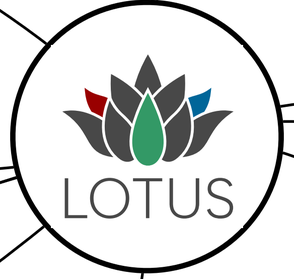
Out today!

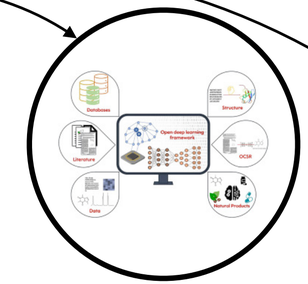


Plant adaptation to heat & drought



Enzyme inhibition with marine NPs



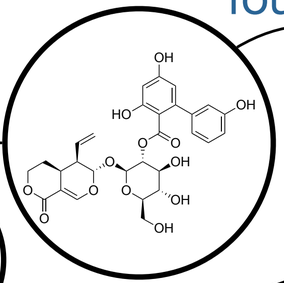


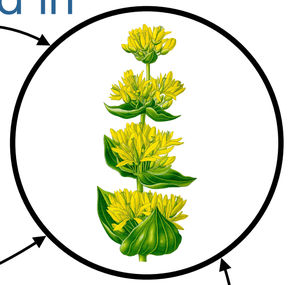
Natural Products & Machine Learning

found in

Bitter ?

?

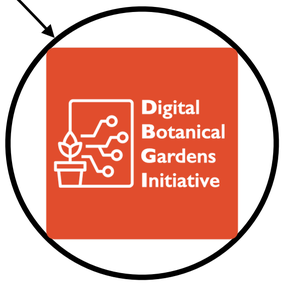




has parts

Next ?

?



Keep it for happy hour!

Acknowledgements

UNIVERSITÉ DE GENÈVE
FACULTY OF SCIENCE
Section of Pharmaceutical Sciences

LIFE SCIENCES PhD SCHOOL GENEVA

Swiss National Science Foundation

WIKIDATA

SCHWABE
Dr. Willmar Schwabe
From Nature. For Health.

GA DUBLIN, IRELAND 2-5 JULY 2023

ETH zürich
D BIOL
Institute of Molecular Systems Biology

Prof. Nicola Zamboni

Contact

rutz@imsb.biol.ethz.ch