Stable identifier initiative, report and discussion of next steps

Anton Güntsch

Freie Universität Berlin

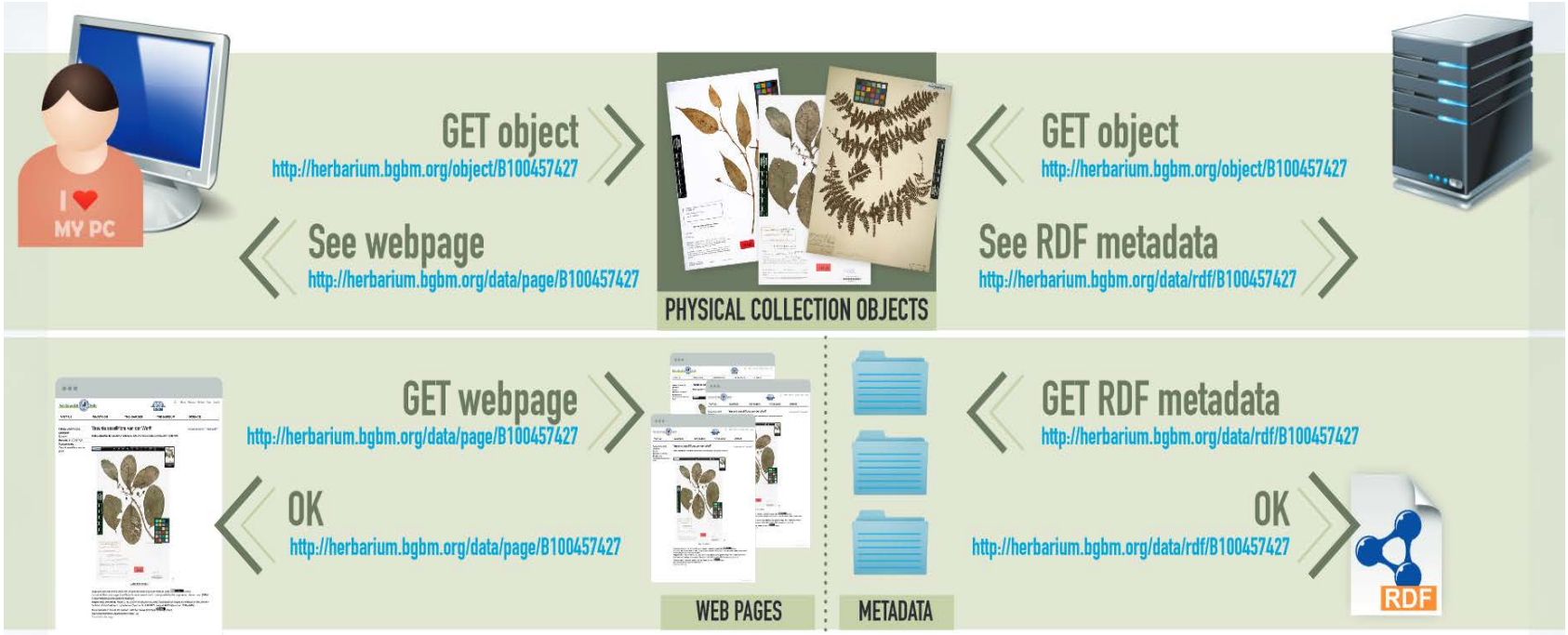
Botanic Garden & Botanical Museum Berlin-Dahlem
Research and Development Group Biodiversity Informatics

BiodiversityInformatics@bgbm.org

For newcomers: what is it all about?

- A joint system of identifiers for specimen held by CETAF collections.
- Use of HTTP-URIs based on Linked Open Data (LOD) principles.
- Redirection mechanism to human-readable webpages and (preferably RDF) Metadata records.

→ Part of the CETAF strategic development plan for the next decade.



The Geneva meeting (10/2016)

Scope of the meeting: to agree on clear implementation steps to be taken by participating partners until the spring 2016 CETAF general meeting and ISTC meeting.

- Berlin (BGBM)
- Berlin (MfN)
- Edinburgh
- Paris
- Geneva
- Vienna
- Stockholm

Priorities for 2015/2016

- To agree on a "CETAF Specimen Preview Profile" (CSPP).
- To implement the CSPP across members of the ISTC identifier initiative.
- To register members of the ISTC identifier initiative together with links to their implementations prominently on the CETAF website.
- To demonstrate the usefulness of the CETAF identifier system using the Wallich Catalogue as a use case.

State of the play (1)

Stable identifier implementations in CETAF member institutions

Botanischer Garten und Botanisches Museum Berlin-Dahlem
Finnish Museum of Natural History, Helsinki
Museum für Naturkunde, Berlin
Muséum national d'histoire naturelle, Paris
Naturalis Biodiversity Center, Leiden
The Natural History Museum, London
Royal Botanic Garden Edinburgh
Royal Botanic Garden Kew
Staatliches Museum für Naturkunde Stuttgart
Zoologisches Forschungsmuseum Alexander Koenig, Bonn

State of the play (2)

Stable identifiers published on institutional websites

Digital Imaging at B
Equipment and process
How to cite this site
Copyright
Barcodes at B

New search [Previous Page / Next Page] Hits: 1

Stable Identifier: http://herbarium.bgbm.org/object/B100457427
Barcode: B 10 0457427

Family: LAURACEAE

Storage name: Yasunia sessiliflora van der Werff

Freie Universität Berlin

Home About us Contact Press Imprint

VISIT US | WHAT'S ON | THE GARDEN | THE MUSEUM | SCIENCE

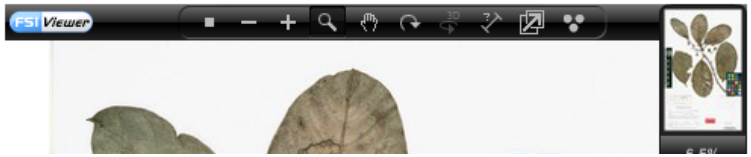
Family: LAURACEAE
Collection: Ecuador.
Barcode: B 10 0457427
Storage name: Yasunia sessiliflora van der Werff
Type status: isotype

Yasunia sessiliflora van der Werff

Annotate specimen User guide

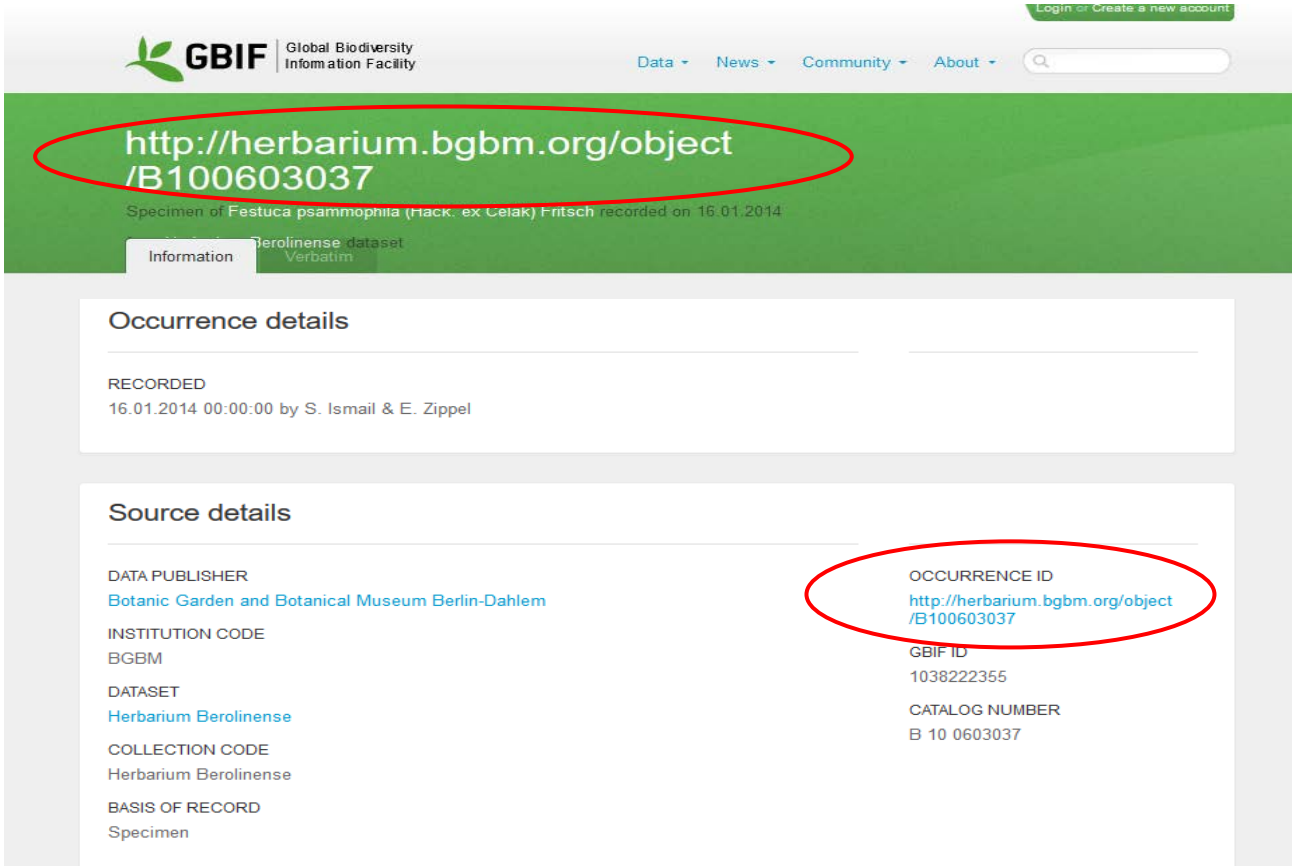
Stable identifier for specimen citations: http://herbarium.bgbm.org/object/B100457427

Save as JPG...



State of the play (3)

Stable identifiers published through GBIF



State of the play (4)

Stable identifiers used in publications

A single collection was cited in the protologue of the name, but none of the three duplicates mentioned above carry the species name nor were they annotated by Hector Léveillé. Lauener (1965) indicated that the type is at E, but he did not specify which of the two sheets there is the type, and therefore a second-step lectotypification is provided here.

3. ***Cardamine circaeoides*** Hook.f. & Thomson, J. Proc. Linn. Soc., Bot. 5: 144. 1861 (http://ipni.org/urn:lsid:ipni.org:names:280232-1:1.1.2.1.1.3). Described from: [INDIA] "In Himalaya orientali temperata, Sikkim interiore, sylvis, alt. 5000-7000 ped. ! J. D. H. (v.v.)". Lectotype (designated here, or perhaps holotype): [Label 1]: [INDIA] "Hab. Sikkim, Regio temp., *J. D. H*[*ooker*]"; [Label 2]: [INDIA] "295 Hab. wet wood, Gohsun, Sikkim, 5000 ft." – K! (K000077050 [http://specimens.kew.org/herbarium/K000077050]); Doubtful isolectotypes – B! (B 10 0386925 [http://herbarium.bgbm.org/object/B100386925]), P! (P00747534 [http://coldb.mnhn.fr/catalognumber/mnhn/p/p00747534]).

The K specimen above was collected at an elevation of 5,000 ft, which is in agreement with the protologue, whereas labels of the B and P sheets indicate the elevation of 6,000–10,000 ft. It is questionable whether the three specimens above were collected from the same area, and that is why we feel that the B and P specimens are doubtful isolectotypes.

PhytoKeys

State of the play (5)

CETAF Specimen Preview Profile agreed.

CSPP Elements

CSPP element	Mandatory	RDF resource URI	Notes
title	yes	http://purl.org/dc/terms/title	Mandatory. Collection may decide themselves on the content of the title element. This could usually be a scientific name (with or without authorship), but it could also contain additional information on collectors and projects for example.
kindOfMaterial	no	http://purl.org/dc/terms/type	The nature of the specimen. Example: herbarium sheet, fossil, conserved in alcohol
scientificNameCurrent	no	http://rs.tdwg.org/dwc/terms/scientificName	The latest full scientific name, with authorship and date information if known.
family	no	http://rs.tdwg.org/dwc/terms/family	The full scientific name of the family in which the taxon is classified.
scientifcNameOriginal	no	http://rs.tdwg.org/dwc/terms/previousIdentifications	The original name appearing on the specimen.
collectorNumber	no	http://rs.tdwg.org/dwc/terms/fieldNumber	An identifier given to the Occurrence at the time it was recorded. Often serves as a link between field notes and an Occurrence record, such as a specimen collector's number.
collectorName	no	http://rs.tdwg.org/dwc/terms/recordedBy	A list (concatenated and separated) of names of people, groups, or organizations responsible for recording the original Occurrence. The primary collector or observer, especially one who applies a personal identifier (recordNumber), should be listed first.
webscaledImageLink	no	http://rs.tdwg.org/dwc/terms/associatedMedia	Link to "nice" webscaled image.
latitude	no	http://rs.tdwg.org/dwc/terms/decimalLatitude	The geographic latitude (WGS84, in decimal degrees, using the spatial reference system given in geodeticDatum) of the geographic center of a Location. Positive values are north of the Equator, negative values are south of it. Legal values lie between -90 and 90, inclusive.
longitude	no	http://rs.tdwg.org/dwc/terms/decimalLongitude	The geographic longitude (WGS84, in decimal degrees, using the spatial reference system given in geodeticDatum) of the geographic center of a Location. Positive values are east of the Greenwich Meridian, negative values are west of it. Legal values lie between -180 and 180, inclusive.
isoCountry	no	http://rs.tdwg.org/dwc/terms/countryCode	The standard (ISO 3166-1 alpha-2) code for the country in which the specimen has been collected. Recommended best practice is to use ISO 3166-1-alpha-2 country codes.
collectionDate	no	http://purl.org/dc/terms/created	(Iso) date of the gathering event.
sourceLink	no	http://purl.org/dc/terms/publisher	URI to a web page with information about the institution/collection/owner responsible for publishing the specimen.

Implemented by RBGE, Paris, BGBM (Kew underway)

State of the play (6)

CETAF URI TESTER | IMPLEMENTERS | DOCUMENTATION | CONTACT

CETAF Specimen URI Tester

Enter a URI in the box below and click 'Check Now' or click for example.

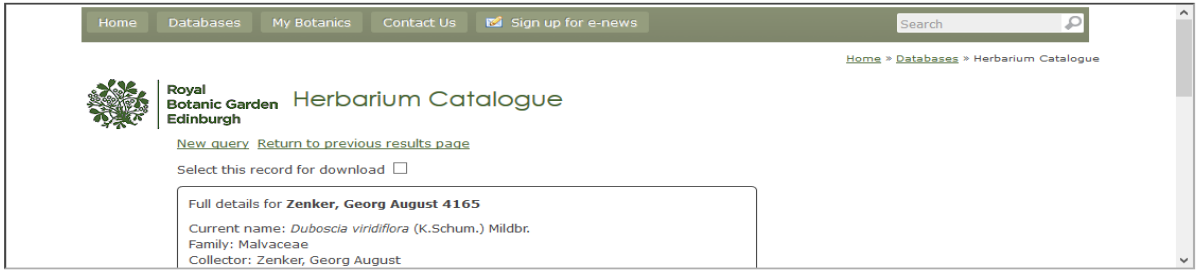
http://data.rbge.org.uk/herb/E00421509 [Check Now]

Checking URI format

✓ The scheme 'http' is correct
✓ The host has the domain name of 'data.rbge.org.uk'
✓ The path component is '/herb/E00421509'
✓ The URI lacks a query string component which is a good thing.
✓ Format of URI appears OK. Continuing test.

Requesting HTML Format Data

ℹ Requesting HTML by passing 'Accept: text/html' header.
✓ Recieved 303 Redirect HTTP code.
✓ Redirect to URI: http://elmer.rbge.org.uk/bgbase/vherb/bgbasevherb.php?cfg=bgbase/vherb/fulldetails.cfg&specimens_specimen__num=469051



ℹ Had 303 redirect when asking for HTML so will request RDF format.

Requesting RDF Format Data

ℹ Requesting RDF by passing 'Accept: application/rdf+xml' header.
✓ Recieved 303 Redirect HTTP code.
✓ Redirect to URI: http://data.rbge.org.uk/service/rdf/herb.php?guid=http://data.rbge.org.uk/herb/E00421509
ℹ Had 303 redirect so will request RDF data for parsing.

Parsing RDF

CETAF Specimen Preview Profile

Prefered URI's are listed first followed by any depricated URIs that may contain the required values.

CSPP Element	Mandatory	RDF resource URI	Resource Expected	Value
title	Yes	**http://purl.org/dc/terms/title**	False	Zenker, Georg August #4165 Duboscia viridiflora (K.Schum.) Mildbr.
Kind of Material	No	**http://purl.org/dc/terms/type**	True	Specimen

State of the play (7)

Discussion / next steps

- Reasonable (**and pragmatic)** next steps?
- Examples:
  - Completion of the Wallish Catalogue Pilot
  - Joint query mechanisms
  - Linking out to other ressources
  - Role out Ids/LOD/CSPP in more CETAF Institutions?
  - Your ideas?
- How? Who?
- The CETAF secretariat could provide support for organising workshops/hackathons/tutorials