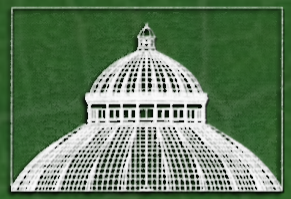


THE NEW YORK BOTANICAL GARDEN

World Flora Online:
Image Metadata Requirements

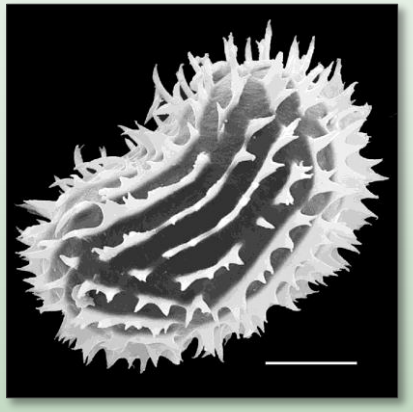
Kimberly Watson ● Information Manager for Digitization

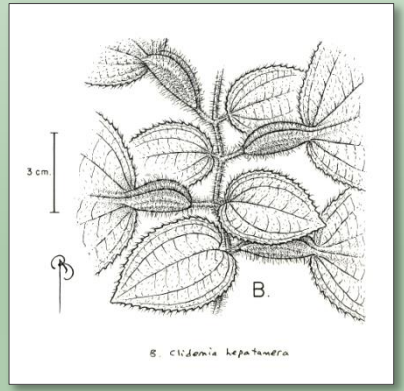
Version: 2015.05.05

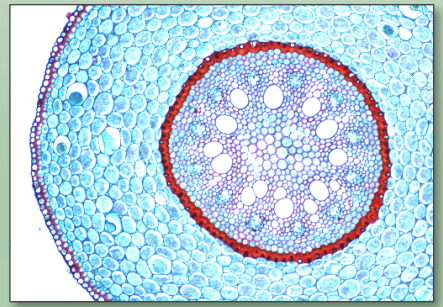
The following presentation was produced to instruct
New York Botanical Garden staff intending to digitize field photographs
for inclusion in the World Flora Online project.
The instructions, software, and examples are specific to NYBG.

Types of Multimedia

1. Still Images
   - Field photos: habitat, habit, key features, etc.
   - SEM and optical microscope micrographs
   - Botanical line drawings
   - Herbarium specimen images

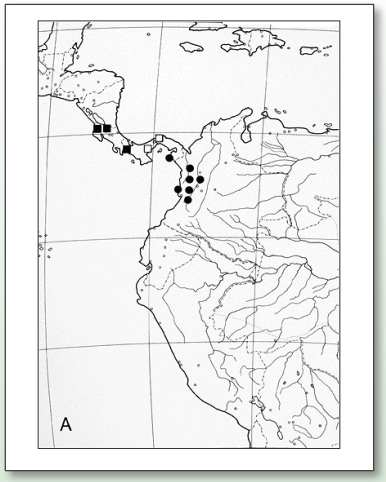


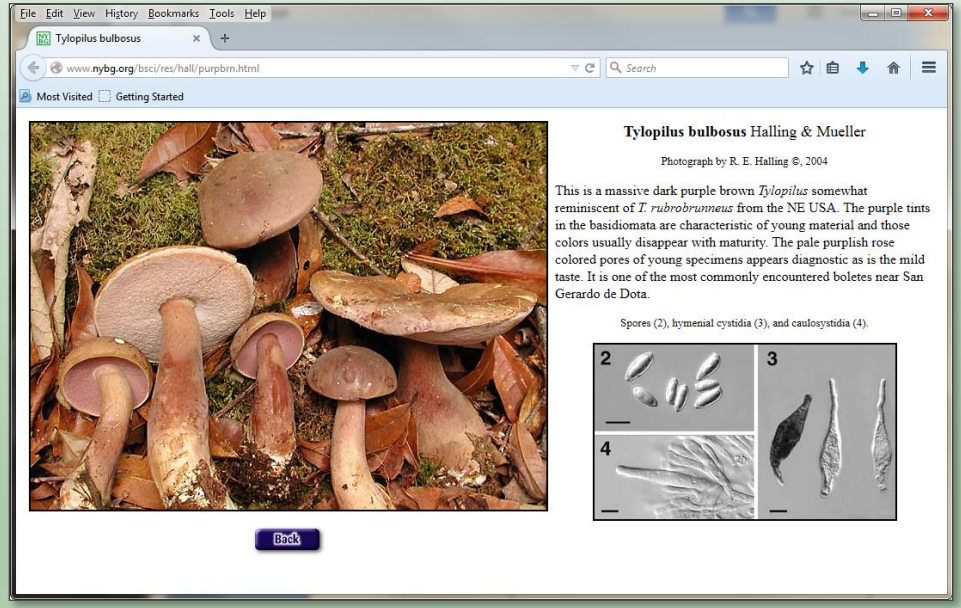


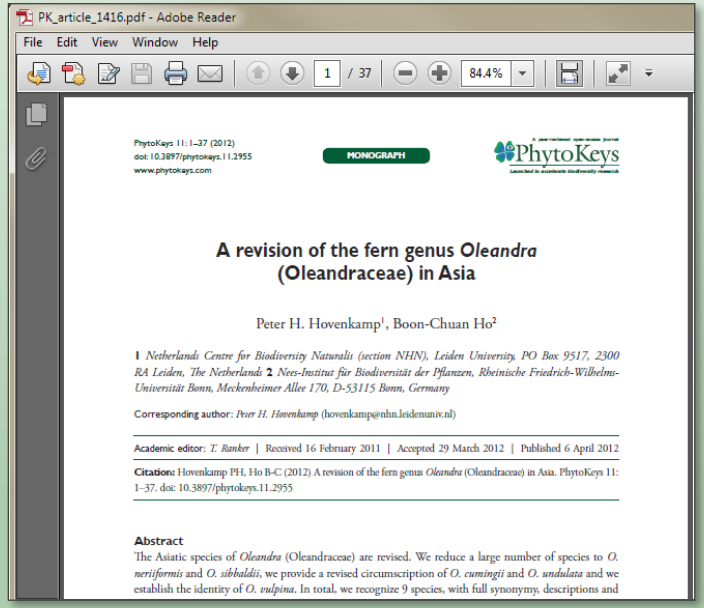


Types of Multimedia

2. Moving Images: MPEG, WMV, AVI, …
3. Sound files: MP3, AIFF, WAV, …
4. Text files: .pdf, .doc, .txt, .csv, .xls, …
5. Interactive Resource: URL for web page







http://www.nybg.org/bsci/res/hall/purpbrn.html

Field Photos

File types for KE EMu:

- **Highest quality JPEG with embedded metadata**
- **If not currently a JPEG, a JPEG copy will be generated with embedded metadata**

File types for archiving:

- **Highest quality original (may be the original JPEG)**
- **If original is camera raw or TIF (including layers), then**
  - **DNG copy will be generated for archive**
  - **Will include embedded metadata**
- **File name = the file name saved in EMu**

Outline

- What is metadata and why it is important
- What metadata is required for WFO
- What metadata is strongly recommended for WFO
- How to capture metadata for your images
  - Using an MS Excel template spreadsheet
  - Using Adobe Lightroom

What is Metadata?

"Structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use, or manage an information resource" (Radebaugh & Guenther 2001).

"Data about Data": 3 Types

Descriptive describe for purposes of discovery and identification (e.g. title, caption/description, keywords, author, etc.)

Structural describe how compound objects are put together (e.g. how pages are ordered to form chapters)

Administrative provide information for managing a resource

- Technical: when/how created, file type, size, color profile, etc.
- Rights management: intellectual property rights, copyright, licensing
- Preservation: information to archive and preserve

Guenther, R. & J. Radebaugh. (2004). Understanding metadata. NISO Press, Bethesda, MD. http://www.niso.org/publications/press/UnderstandingMetadata.pdf

Why is Metadata important?

Can be stored in the image wherever it goes

Can be stored in a catalog or other digital asset management system

- **Organize electronic resources**
- **Facilitate resource discovery**
  - Allow resources to be found by relevant criteria
  - Identifying resources
  - Bring similar resources together
  - Distinguish dissimilar resources
  - Provide location information
- **Provide digital identification**
- **Facilitate interoperability and legacy resource integration**
- **Support archiving and preservation**

Guenther, R. & J. Radebaugh. (2004). Understanding metadata. NISO Press, Bethesda, MD. http://www.niso.org/publications/press/UnderstandingMetadata.pdf

Metadata Schemes:

- **IPTC:** International Press and Telecommunications Council
- **XMP:** Adobe Extensible Metadata Platform
- **EXIF:** Exchangeable Image File Format

- **DC: Dublin Core Metadata Initiative**, defines elements for use by authors to describe their Web-based resources. (http://dublincore.org/documents/dcmi-type-vocabulary/)

- **AC: Audubon Core Multimedia Resources Metadata**, vocabularies designed to represent metadata for biodiversity multimedia resources and collections (http://terms.tdwg.org/wiki/Audubon_Core)
    - Management of the media and collections
    - Descriptions of their content
    - Taxonomic, geographic, and temporal coverage
    - Appropriate ways to retrieve, attribute, and reproduce them

Metadata Required for WFO

Term	Definition
File Name = Identifier	**An arbitrary code that is globally unique for the resource.** **= Unique among all multimedia resources** **= Lowercase letters a-z, 0-9, hyphen, or underscore** **= No spaces, punctuation, symbols, or accents** **Examples:** kwatson_37.jpg cavendishia_callista2.jpg Clidemia_octona_Ventura_12898_Mexico_CAS_04_Gilberto_Ocampo.jpg **NOTE: Please discuss with Charlie Zimmerman how to proceed, if…** • current file names do not meet these criteria • and/or they contain human-readable information that could be parsed into other metadata fields such as scientific name, creator, collector and number, etc. **Example:** Dis stereophyllum Salinas-663 (Betancur)-0002, Chocó-Valle.jpg

Metadata Required for WFO

Term	Definition
Scientific Name	**Full scientific name (genus species author) as it appears in IPNI or The Plant List** (http://www.theplantlist.org/) **Example:** Acanthopale confertiflora (Lindau) C.B.Clarke
Creator	**The full name of the person or organization responsible for creating the media resource (i.e. photographer).** Name will be matched with an existing Parties record in KE EMu, or will be the basis for a new Parties record.

Metadata Required for WFO

Term	Definition
Copyright Statement	**Information about rights held in and over the resource.** **A full-text, readable copyright statement, as required by the national legislation of the copyright holder.** **Common example:** "Rights reside with creator, otherwise property of NYBG." "Data (images and text) contained in the Herbarium Specimen Catalogs are copyrighted by The New York Botanical Garden. Use or reproduction of these data is authorized for educational or other non-commercial purposes without prior permission from the copyright holders, provided credit is given to The New York Botanical Garden. Use or reproduction of these data for commercial purposes is prohibited without prior written permission of The New York Botanical Garden." **Creative Commons Licensing:** http://creativecommons.org/licenses/ CC BY NC SA **Considerations for licensors and licensees:** https://wiki.creativecommons.org/Considerations_for_licensors_and_licensees

Metadata Recommended for WFO

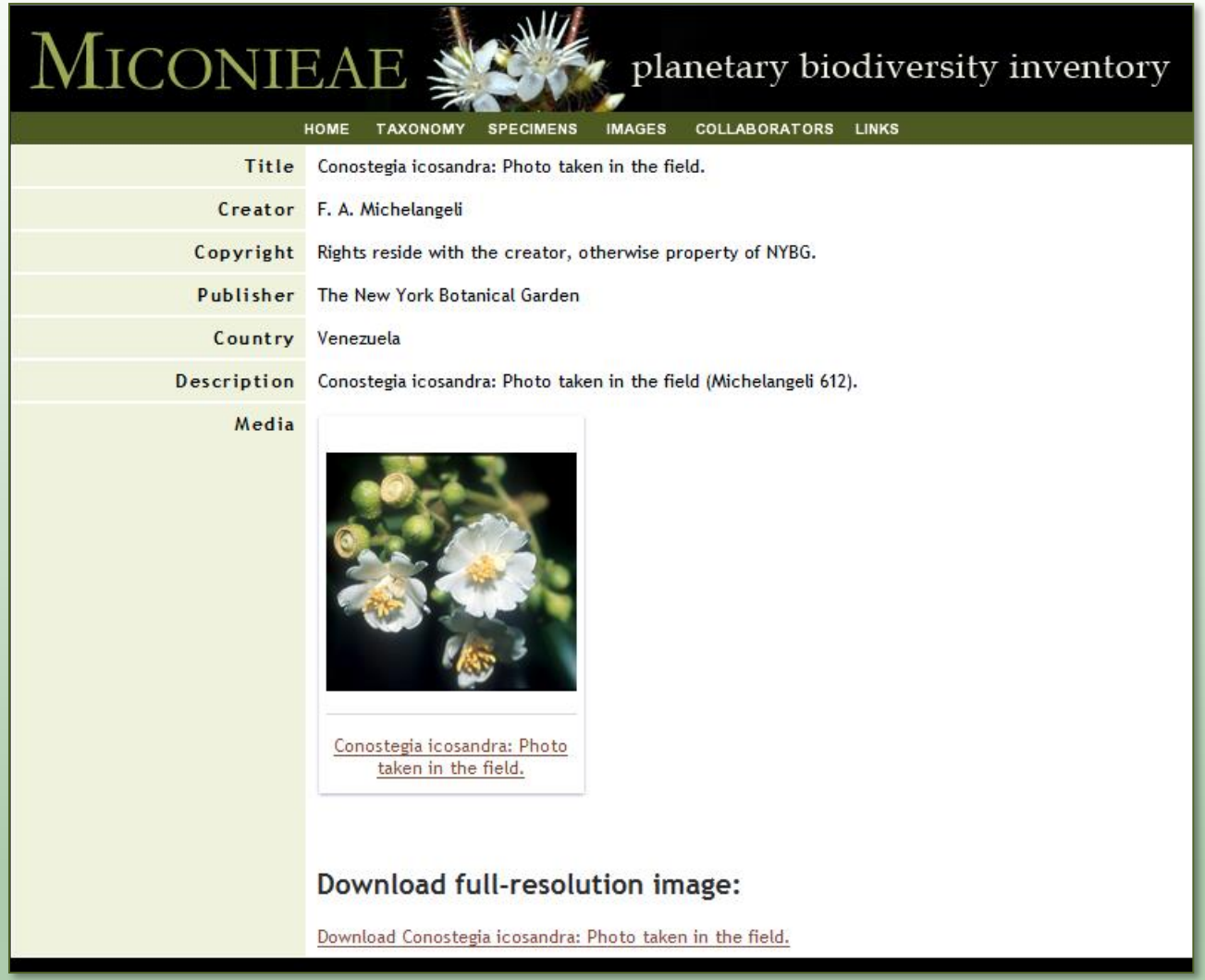
Term	Definition
Title	**Concise title, name, or brief descriptive label of the individual resource.** This facilitates interactions with humans, e.g. as display text of a hyperlink, or as a choice in a pick-list of images. Accents and punctuation are accepted. It is not recommended to include the scientific name, as this may change over time. **Example:** Field photo of Fabián Michelangeli 1587.
Description	**(= Caption) An account or description of the content of the resource written as free-form text.** Ideally this will include the Who, What, Where, When, Why of the resource. May include but is not limited to: an abstract, a table of contents, a graphical representation, or a free-text account of the resource. Accents and punctuation are accepted. **Example:** Photograph taken in the field of Michelangeli 1587 (NY 02103779), West Indies, Cuba, Holguín, Moa, Subida a Santa Teresita, 2 km E de Yamaniguey y de alli subida hacia el S-SW 3.5-5.5 km.

Metadata Recommended for WFO

Term	Definition
Collector	**Full name of the primary collector who collected the specimen in the photograph, if known.** Collector's name will be matched with an existing Parties record in KE EMu, or will be the basis for a new Parties record.
Collection Number	**The number assigned to the collection by the collector, if known.**
Country	**The geographic location of the specific object(s) documented in the resource.** For example, the country in which the plant was growing when photographed in the field.
Keywords	**General keywords or tags.** May be multi-worded phrases. No need to include scientific name, country, or other metadata elements captured in more specific fields. **Examples:** Plant parts photographed, habitat type, plant habit, life stage, plant family, etc.

Sample Mini-Site

PBI: Miconieae website demonstrates how image metadata could be displayed online.



How to Capture Metadata for WFO

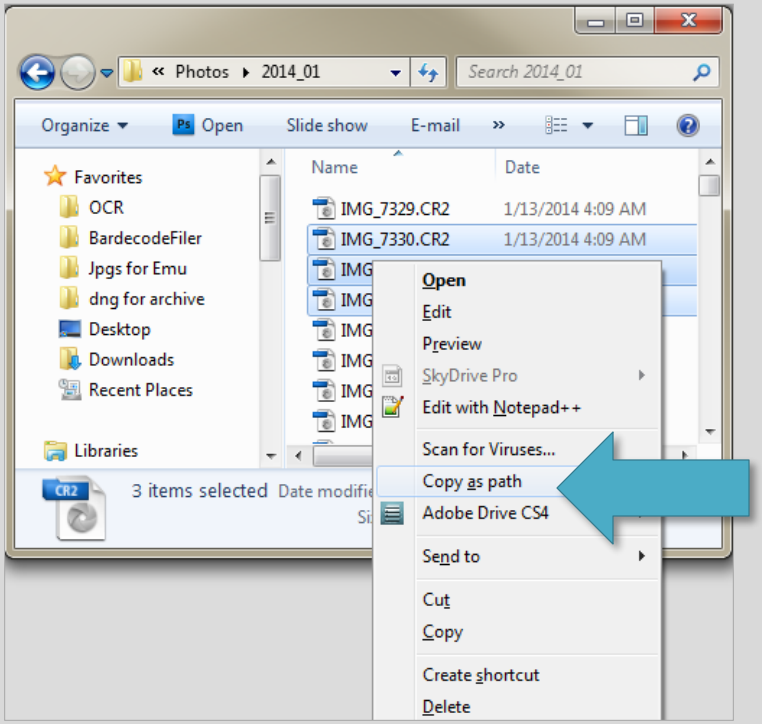
Option 1: **Use the Excel template spreadsheet saved here:**
X:\Pub\World Flora Online\WFO_Image_Metadata_Template_Mar2015.xlsx

	A	B	C	D	E	F
2	**Field**	**File Path**	**Filename = identifier**	**Creator**	**Scientific Name**	**Plant Family**
3	**Format**	REQUIRED. The full path to the resource	REQUIRED. Any multimedia file with extension: pdf, jpg, txt, doc, xls, etc.	REQUIRED. Text. Name of photographer or source of multimedia resource (preferably full name: first name, middle initial, last name. Could be name of an organization)	REQUIRED. Text. Scientific name of the plant captured in the image, including authors.	Text.
4	**Example**	C:\Field_Photos\EschAguilarii.jpg	EschAguilarii.jpg	Scott A. Mori	Eschweilera longirachis S.A.Mori	Lecythidaceae
5	**Start here →**					
6						
7						

If choosing this option, additional image metadata is required, as defined on the following 2 slides.

Metadata Required for Excel Spreadsheet

Term	Definition
Full Path	Specifies the root directory and all other subdirectories that contain the resource. Must be less than 250 characters in length **Example:** C:\Watson\Field_photos\2011-06-23\Ossaea_micrantha\kwatson_37.jpg



How to "Copy as Path":
Select images
Hold Shift and right-click mouse
Select "Copy as path"
Paste file path(s) into Excel spreadsheet

"H:\Photos\2014_01\IMG_7331.CR2"
"H:\Photos\2014_01\IMG_7332.CR2"
"H:\Photos\2014_01\IMG_7330.CR2"

Metadata Required for Excel Spreadsheet

Term	Definition	Controlled Vocabulary
Type	**Any type from the DCMI Type Vocabulary** (http://dublincore.org/documents/dcmi-type-vocabulary/#H7) Choose one option from the controlled list on the right.	**StillImage** Sound MovingImage InteractiveResource Text Collection
Subtype	**Any subtype from this list.** **NOTE:** If a resource cannot be classified as any of these subtypes, please contact Charlie Zimmerman for guidance.	**Field photograph** Handwriting sample Herbarium specimen Label image SEM Drawing
Metadata Language	**Language of the metadata (title, description, etc.),** not necessarily of the image. Represented as ISO639-2 three-letter code (http://id.loc.gov/vocabulary/iso639-2.html)	**eng** **spa** **fre** **por**

How to Capture Metadata for WFO

Option 2: **Use Adobe Lightroom to organize images and capture metadata**