

Imaging Collections

Loïc Costeur



Why?

support management

facilitate everyday work

facilitate asset identification

facilitate digitization

open new projects

improve long-term conservation

serve research purposes

serve educational/promotional purposes

The image displays a digital collection management system interface. On the left, a spreadsheet lists specimen data with columns for 'Standort' (Location), 'Fundort' (Findspot), and 'MP-M' (Museum Project). The data includes various locations such as Orsmael, Brabant = Dormaal, Harwich (GB), Herne Bay (GB), Londonthor (GB), Pourcy, Marne (FR), Meudon, Seine-et-Oise (FR), Palette bei Aix-en-Provence (FR), Condé-en-Brie, Aisne (FR), Avenay, Marne (FR), and Monthelon, Marne (FR). The 'MP-M' column lists museum projects like MP7, MP8-9, MP8-9?, MP10, and MP10?.

In the center, a window titled 'K4_201807_Monthelon.JPG - IrfanView' shows a photograph of a specimen tray. The tray is filled with numerous small, labeled compartments, each containing a specimen, likely insect remains, with handwritten labels.

On the right, a grid of thumbnail images shows various specimen trays and boxes, each labeled with a unique identifier such as 'K4_201A07_Biberfeld', 'K4_201A08_Shepton-Mallet', 'K4_201A09_Peru-Sillustani', 'K4_201A10_Cha mp-Garimond_Fo ns', 'K4_201A16_Cern ay-les-Reims', 'K4_201A17_Muti gny', 'K4_201A19_Berru', 'K4_201A20_Berru', 'K4_201A29_Orsm ael', 'K4_201A30_Orsm ael', 'K4_201A31_Orsm ael', 'K4_201A32_Hawi ch_Herne-Bay_Lo ndon-Clay', 'K4_201A38_Aven ay', 'K4_201807_Mont helon', 'K4_201809_Mont helon', 'K4_201811_Mont helon', 'K4_201B22_Mont helon', 'K4_201B24_Mont helon', 'K4_201B26_Mont helon', 'K4_201B28_Mont helon', 'K4_201B36_La-Fé re', 'K4_201B38_Chav ot-Courcut', 'K4_201C01_Belge -Balegeni-Dalun- Osteroden_Merze n', 'K4_201C02_Vertai n-bei-Epernay', 'K4_201C09_Cuis- bei-Epernay', 'K4_201C10_Cuis- bei-Epernay', 'K4_201C11_Manc y', 'K4_201C13_Manc y', 'K4_201C17_Egerk ingen-Marsupiali', 'K4_201C19_Egerk ingen-Insectivora', 'K4_201C20_Egerk ingen-Chiroptera', 'K4_201C21_Egerk ingen-Amphi-He', 'K4_201C22_Egerk ingen-Adapis', and 'K4_201C23_Egerk ingen-Caenophi'.

What?

Collection units

Drawer

Box

Labels

Objects

Image



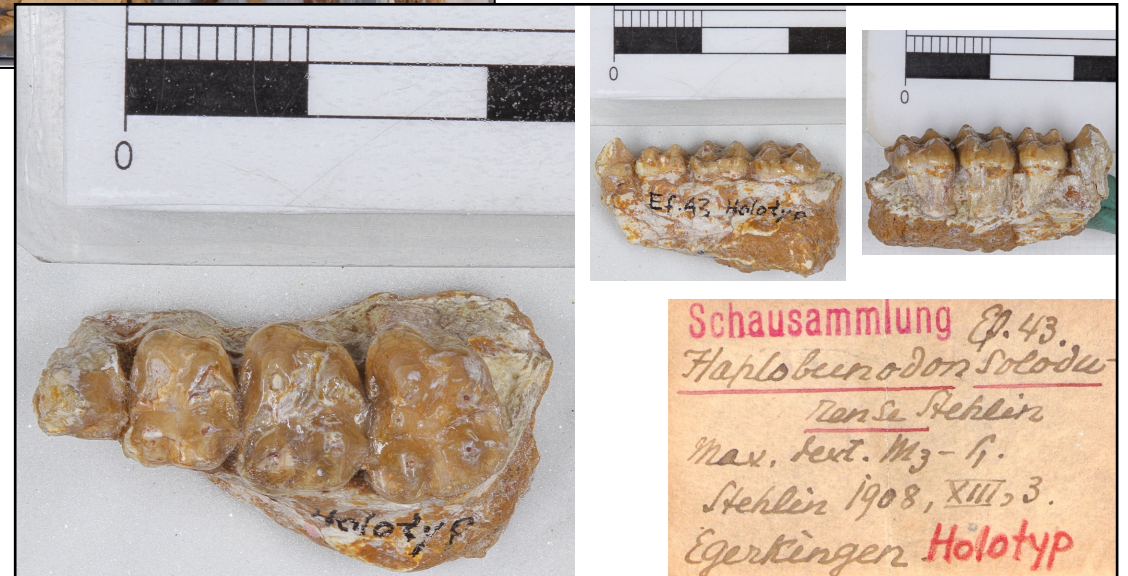
easy

Time
consuming



“Drawer-level digitization has become the most practical way of unlocking the research potential for natural history collections”

Hudson et al., 2015 – PLoS ONE



How?

How you want and how you can !!

Be pragmatic

Depends on:

PURPOSE

collection size & type

2D

3D

external/internal

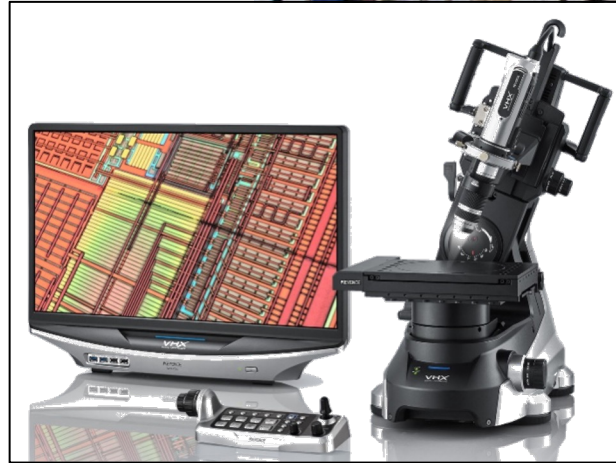
unit size

time

resources



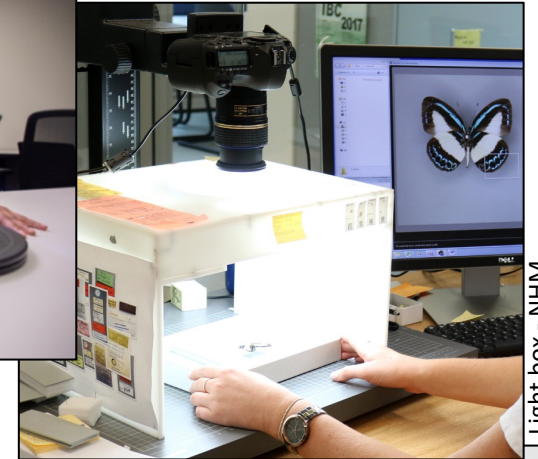
Conveyor Belt - Picturae



Digital microscope - Keyence



3D surface scanner - Artec



Light box - NHM



CT-scanner - GE



Archivscanner-Museumsscanner - Walter Nagel

How?

How you want and how you can

Be pragmatic

Depends on:

PURPOSE

collection size & type

2D

3D

external/internal

unit size

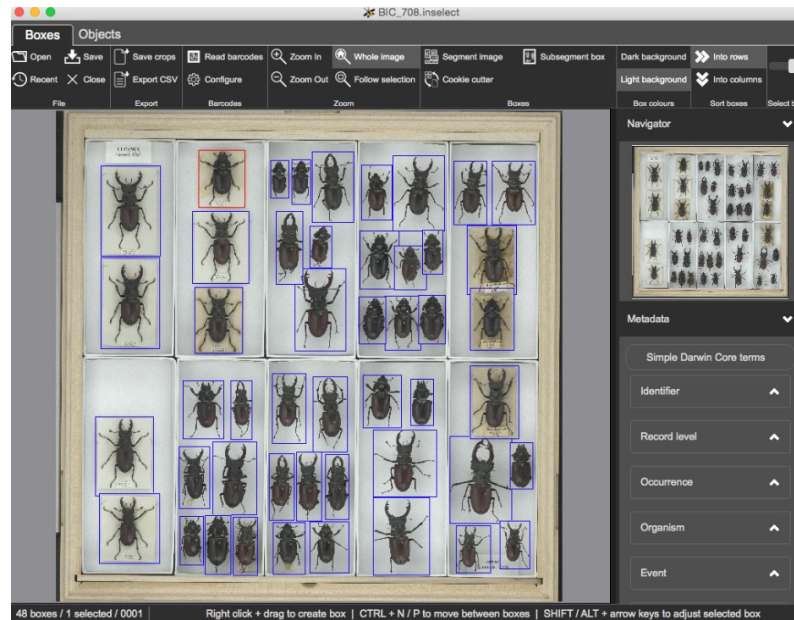
time

resources



Inselect

Open source



PLOS ONE

OPEN ACCESS PEER-REVIEWED

RESEARCH ARTICLE

Inselect: Automating the Digitization of Natural History Collections

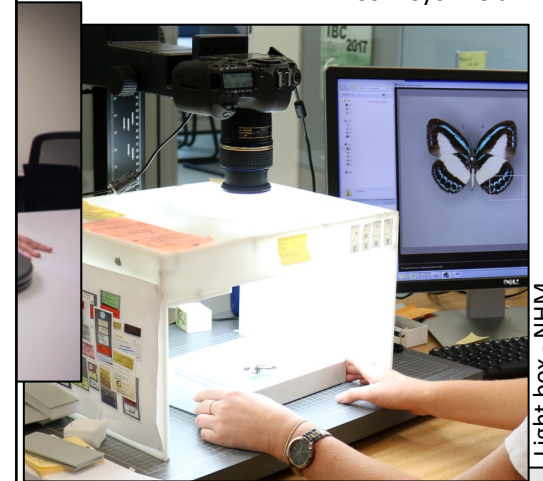
Lawrence N. Hudson, Vladimir Blagoderov, Alice Heaton, Pieter Holtzhausen, Laurence Livermore, Benjamin W. Price, Stéfán van der Walt, Vincent S. Smith

Published: November 23, 2015 • <https://doi.org/10.1371/journal.pone.0143402>

<https://naturalhistorymuseum.github.io/inselect/>



Conveyor Belt - Picturae



Light box - NHM



Archivscanner-Museumsscanner - Walter Nagel

Technical recommendations

For 2D static images

at least 300 ppi
RAW or TIFF (JPEG)

capture data (don't do art)
stack for depth when possible

image labels (600 ppi)

homogeneous background
(adapted to object colour)

scalebar / colour chart

light box when possible
(Colour Rendering Index >95)

Licence CC/Rights holder

For whole-drawer imaging

try bar-coding the drawers
(before)

try homogeneous pictures

For 3D images

OBJ or PLY (STL)

DICOMs (TIFF) for CT

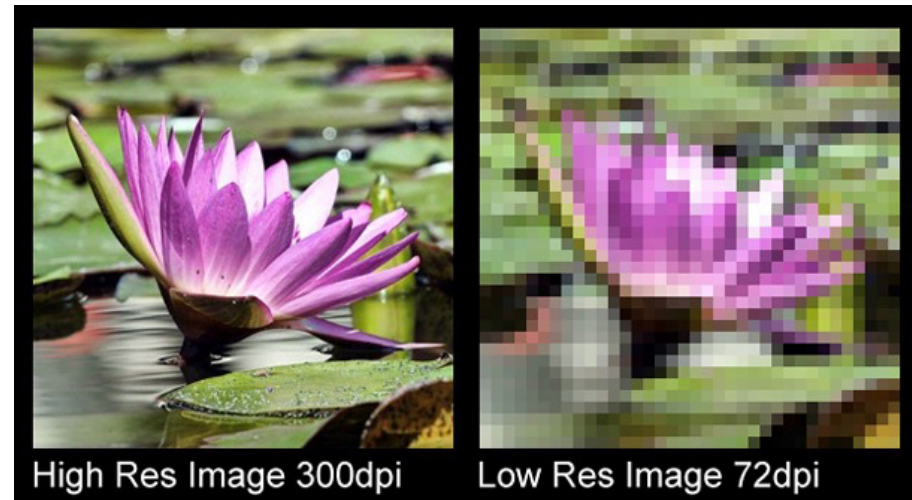
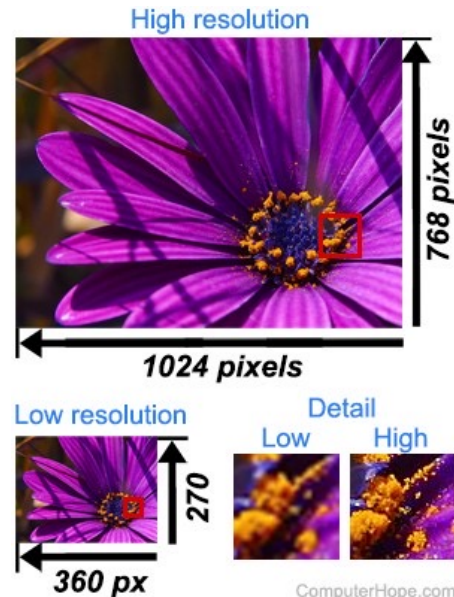
record scanning resolution
(add info file as attachment)

For videos & audios

MOV or MPEG-4 (h.265)

MP3 or WebM

Licence CC/Rights holder



Database fields

Just a proposal – feel free to do as you want!

Synthesis from DarwinCore – AudubonCore – DublinCore – ABCD

class	MachineObservation / Multimedia Object				
class	Image Object	StillImage	MovingImage	Sound	Text
definition	visual representation	static image representation	visual representation with motion	resource to be heard	image with text
properties	title				
	creator				
	date	<i>(e.g., 2022-01-02T12:56:33.000+00:00)</i>			
	licence	<i>(e.g., http://creativecommons.org/...)</i>			
	accessrights	<i>(e.g., not-for-profit use only)</i>			
	rights holder / suggested citation				
	capture equipment	<i>(e.g., canon EOS 5D, lens EF 50mm / GE Phoenix nanotom / Artec 3D space spider)</i>			
	url				
	height (px)	height (px)	height (px)	duration time (s)	
	width (px)	width (px)	width (px)		
	image resolution (ppi)	image resolution (ppi)			image resolution (ppi)
	file size (MB)				
	file format				
	OBJ, PLY, STL...	RAW, TIFF, JPEG...	MPG, MOV, AVI...	MP3, WAV, WMA...	RAW, TIFF, JPEG

Database fields

Image Object Class

Class to describe an image.

i More information

Associated Properties:

width

height

imageResolution

colorSpace

lensModel

lightSource

color

thumbnailURL

fNumber

focalLength

focalLengthIn35mmFilm

flash

flashEnergy

whiteBalance

digitalZoomRatio

contrast

saturation

sharpness

gamma

exposureTime

exposureMode

spectralSensitivity

photographicSensitivity

isoSpeed

shutterSpeed

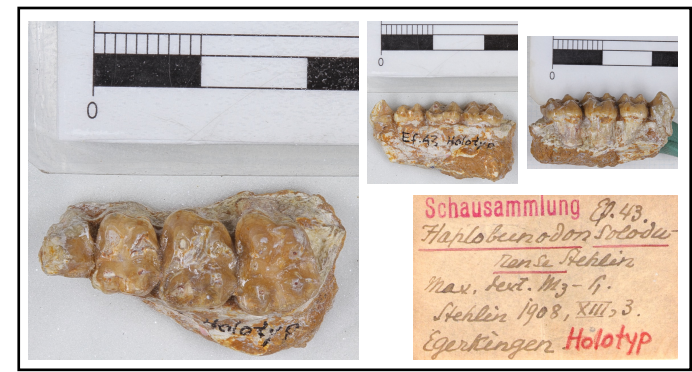
brightness

class					
class	Image Object				
definition	visual representation				
properties					
	date	(e.g., 2012-01-01)			
	licence	(e.g., CC BY-NC-SA)			
	accessrights	(e.g., CC BY-NC-SA)			
	rights holder / suggested citation				
	capture equipment (e.g., canon EOS 5D, lens EF 50mm / GE Phoenix nanotom / Artec 3D space spider)				
	url				
	height (px)	height (px)	height (px)	duration time (s)	
	width (px)	width (px)	width (px)		
	image resolution (ppi)	image resolution (ppi)			image resolution (ppi)
	file size (MB)				
	file format				
OBJ, PLY, STL...	RAW, TIFF, JPEG...	MPG, MOV, AVI...	MP3, WAV, WMA...	RAW, TIFF, JPEG	

Case study

Imaging reference fossils

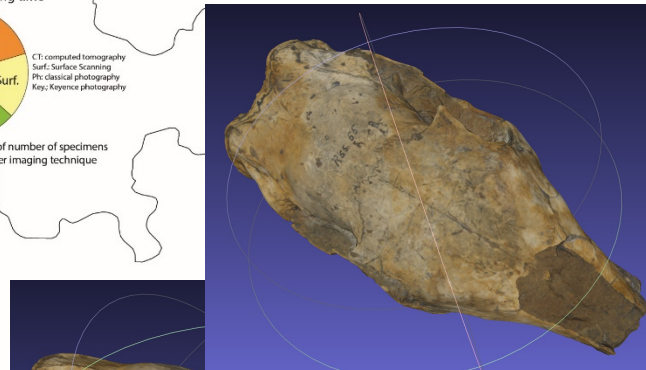
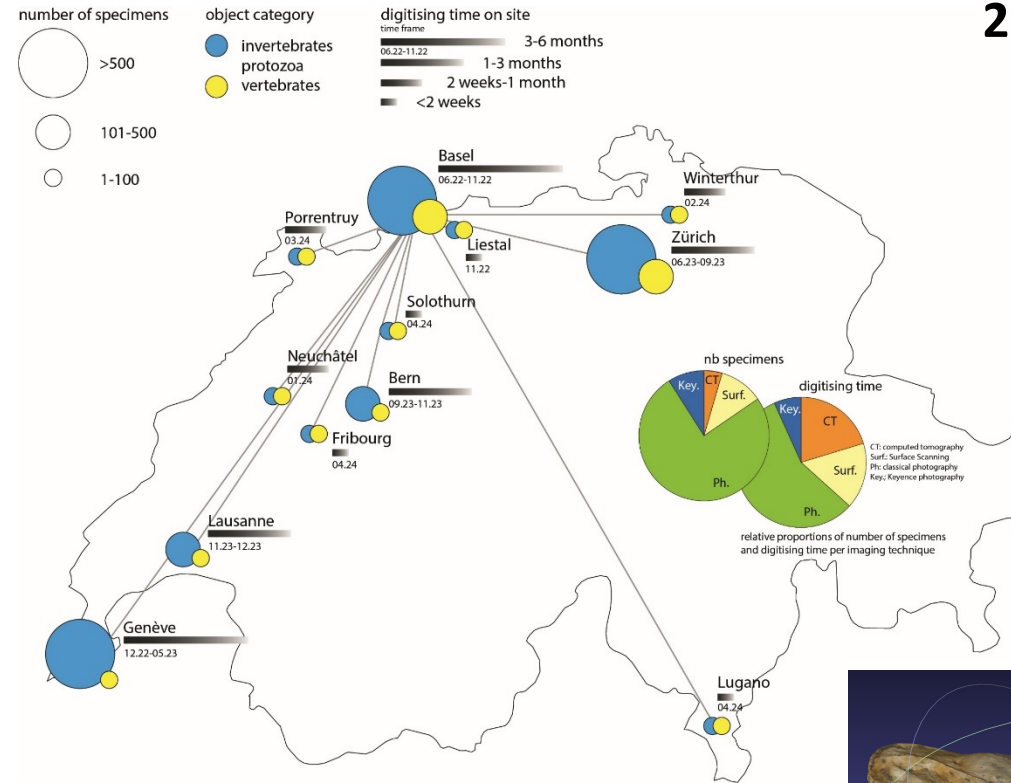
About 10'000 specimens
12 institutions
2 years



NMB Ef.43

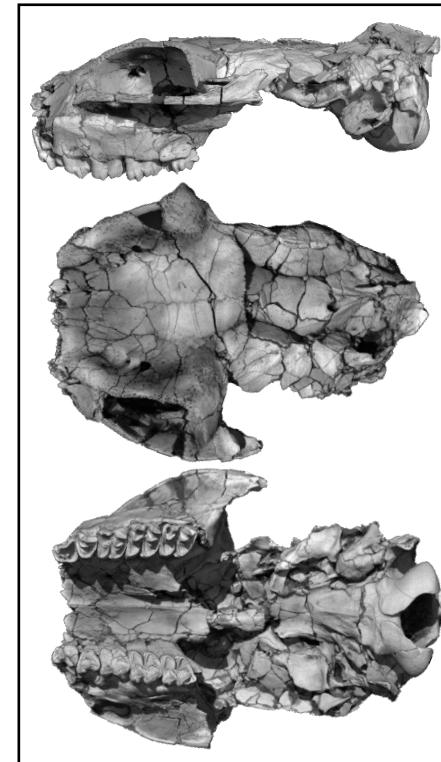
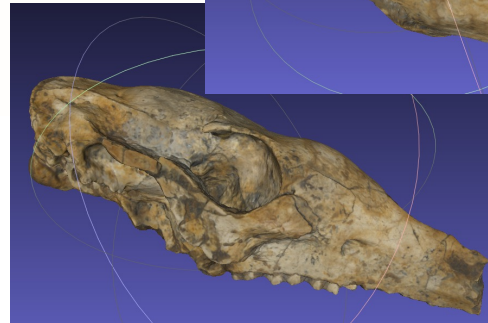
HOLOTYPE

dig. microscope



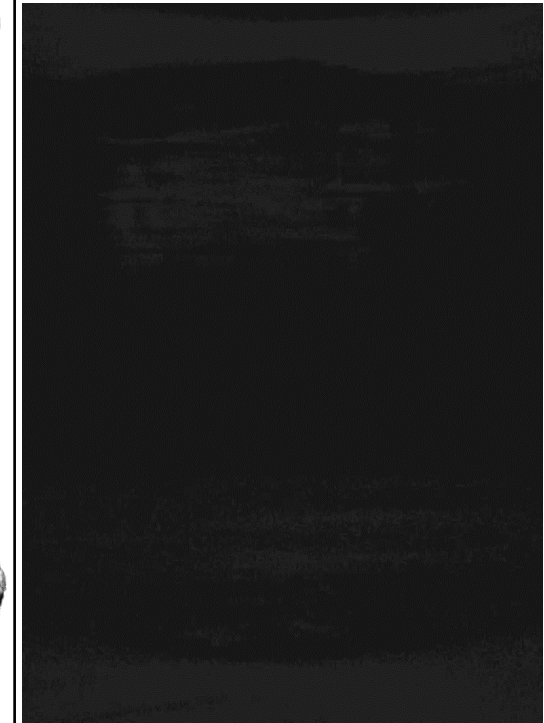
NMB Rss.55

HOLOTYPE
surface scan



NMB Sth.833

PUBLISHED
CT-scan



Difficulties

TIME

Mass production

Very heterogeneous objects

4 different techniques

Databasing

(when data not yet available –
taxonomy-publication records etc)

STORAGE

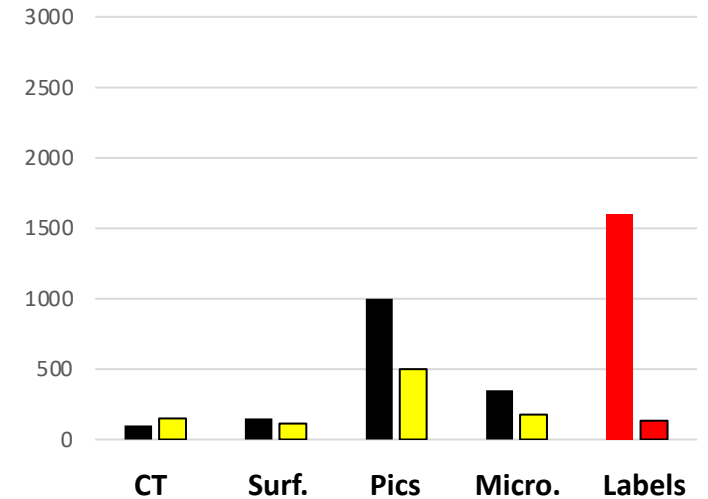
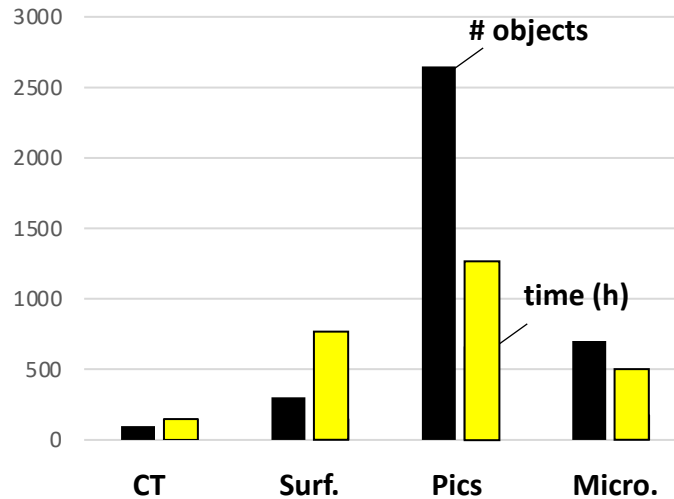
Average 1.55 GB / object

350 objects

550 GB

81'500 files

Data for NMB



Recommendations

be pragmatic & flexible

e.g., be happy with 1 good image (reduce to 1 view)

be less optimistic for time as you usually are 😊

e.g., count a total of at least 30 minutes per object

book 20% of your time for databasing

References



Access Right

<https://www.fieldmuseum.org/field-museum-natural-history-conditions-and-suggested-norms-use-collections-data-and-images>

Current standards



Audubon Core
Multimedia Resources
Metadata Schema

<https://ac.tdwg.org/termlist/>



Darwin Core

Taxonomic Database Working Group – Biodiversity Information Standards

<https://www.tdwg.org/standards/>

<https://www.museumbund.de/wp-content/uploads/2022/12/handreichung-digitale-grunderfassung.pdf>



<https://abcd.tdwg.org/terms/#group-MultimediaObject>

ABCD

Access to Biological
Collection Data

Multimedia Object