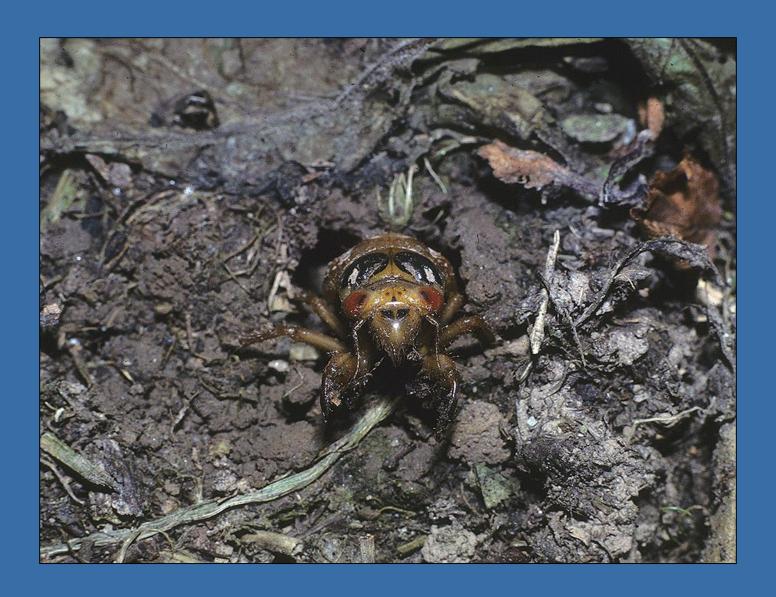
## Workflows: Perspectives from Peabody Entomology



Larry Gall
Systems Office
Yale Peabody Museum

## "On The Ground, In-The-Trenches" View



## "On The Ground, In-The-Trenches" View



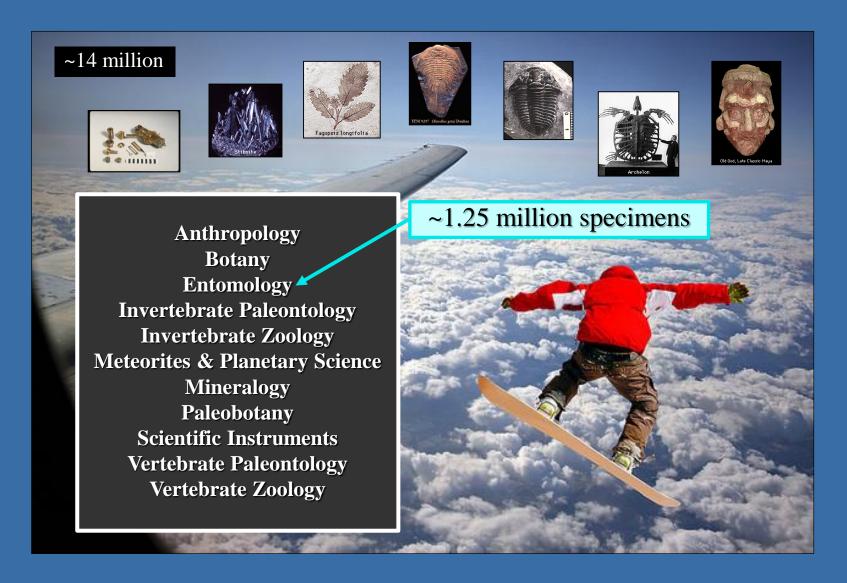


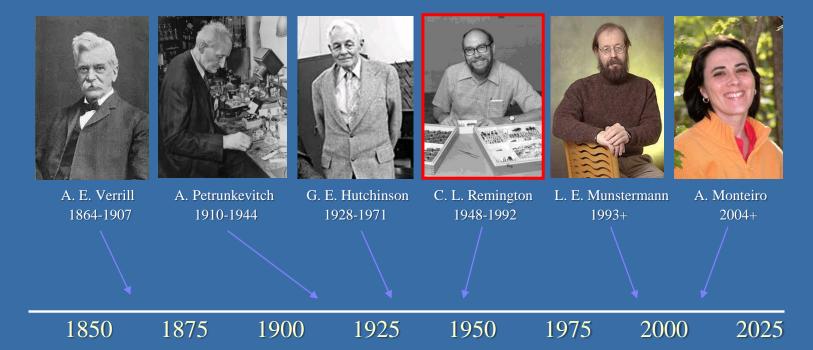
### Institutional History, Scale & Contexts - Strongly Influence Digitization

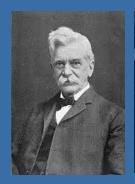
















A. Petrunkevitch 1910-1944



G. E. Hutchinson 1928-1971



C. L. Remington 1948-1992



L. E. Munstermann 1993+



A. Monteiro 2004+

1850

1875

1900

1925

1950

1975

2000

2025

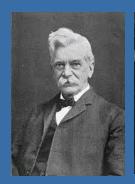


Peabody Museum 1865+





Kline / Osborn Labs 1960-2009



A. E. Verrill 1864-1907



A. Petrunkevitch 1910-1944



G. E. Hutchinson 1928-1971



C. L. Remington 1948-1992



L. E. Munstermann 1993+



A. Monteiro 2004+

1850

1875

1900

1925

1950

1975

2000

2025



Peabody Museum 1865+

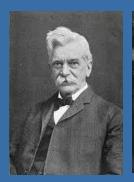




Kline / Osborn Labs 1960-2009



Environmental Science Center (ESC) 2001



A. E. Verrill 1864-1907



A. Petrunkevitch 1910-1944



G. E. Hutchinson 1928-1971



C. L. Remington 1948-1992



L. E. Munstermann 1993+



A. Monteiro 2004+

1850

1875

1900

1925

1950

1975

2000

2025



Peabody Museum 1865+





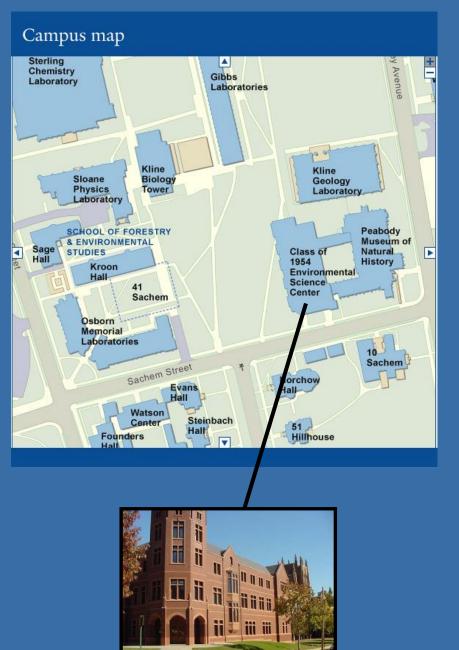
Kline / Osborn Labs 1960-2009



Environmental Science Center (ESC) 2001

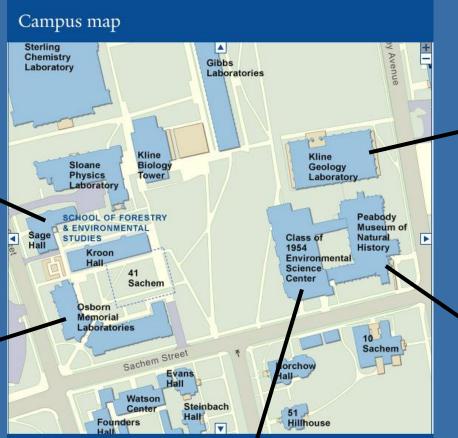


Yale West Campus 2009

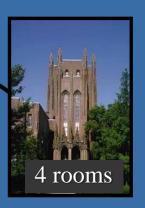




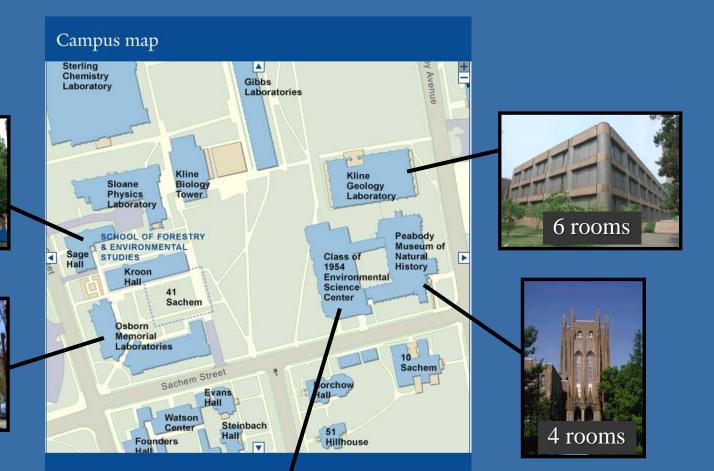








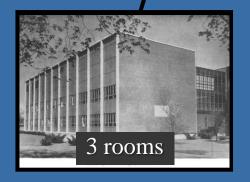


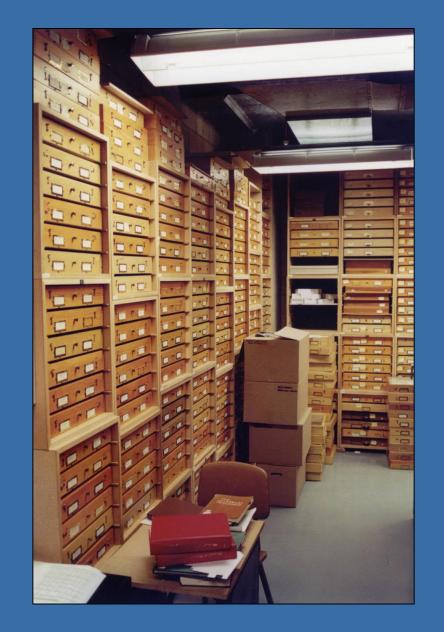


Over 30% growth in size of Entomology collection toward the end of the Remington era

1 room

3 rooms





























# "The "Laying On of Hands"











## "The "Laying On of Hands"











1986 - The Cretaceous of Computing



1986 - The Cretaceous of Computing





• 1986 Most Divisions starting to use personal computers





- 1986 Most Divisions starting to use personal computers
- 1989 Computer Committee studies long-term solutions









- 1989 Computer Committee studies long-term solutions
- 1991 Systems Office created & staffed
- 1991 Argus collections management initiative starts









- 1989 Computer Committee studies long-term solutions
- 1991 Systems Office created & staffed
- 1991 Argus collections management initiative starts
- 1994 Divisions heavily engaged in electronic cataloguing





- 1986 Most Divisions starting to use personal computers
- 1989 Computer Committee studies long-term solutions
- 1991 Systems Office created & staffed
- 1991 Argus collections management initiative starts
- 1994 Divisions heavily engaged in electronic cataloguing
- 1994 Gopher services launched for collections data









- 1986 Most Divisions starting to use personal computers
- 1989 Computer Committee studies long-term solutions
- 1991 Systems Office created & staffed
- 1991 Argus collections management initiative starts
- 1994 Divisions heavily engaged in electronic cataloguing
- 1994 Gopher services launched for collections data
- 1998 Physical move of collections to ESC "begins"









# Peabody Collections

### Digital History & Timeline



- 1986 Most Divisions starting to use personal computers
- 1989 Computer Committee studies long-term solutions
- 1991 Systems Office created & staffed
- 1991 Argus collections management initiative starts
- 1994 Divisions heavily engaged in electronic cataloguing
- 1994 Gopher services launched for collections data
- 1998 Physical move of collections to ESC "begins"
- 2002 Physical move of collections to ESC "ends"









### Digital History & Timeline



- 1986 Most Divisions starting to use personal computers
- 1989 Computer Committee studies long-term solutions
- 1991 Systems Office created & staffed
- 1991 Argus collections management initiative starts
- 1994 Divisions heavily engaged in electronic cataloguing
- 1994 Gopher services launched for collections data
- 1998 Physical move of collections to ESC "begins"
- 2002 Physical move of collections to ESC "ends"
- 2003 Digitization Committee searches for Argus successor









# Peabody Collections

#### Digital History & Timeline



- 1986 Most Divisions starting to use personal computers
- 1989 Computer Committee studies long-term solutions
- 1991 Systems Office created & staffed
- 1991 Argus collections management initiative starts
- 1994 Divisions heavily engaged in electronic cataloguing
- 1994 Gopher services launched for collections data
- 1998 Physical move of collections to ESC "begins"
- 2002 Physical move of collections to ESC "ends"
- 2003 Digitization Committee searches for Argus successor
- 2004 KE EMu selected, preparation for new system begins











# Peabody Collections

#### Digital History & Timeline



• 1986 Most Divisions starting to use personal computers







- 1994 Divisions heavily engaged in electronic cataloguing
- 1994 Gopher services launched for collections data
- 1998 Physical move of collections to ESC "begins"
- 2002 Physical move of collections to ESC "ends"
- 2003 Digitization Committee searches for Argus successor
- 2004 KE EMu selected, preparation for new system begins
- 2005 Argus data migration and go-live in KE EMu













• 2005 Collections data launched to GBIF, Paleoportal et al.





- 2005 Collections data launched to GBIF, Paleoportal et al.
- 2006 Digital conversion of transaction records accelerates
- 2006 Digital imaging of specimens rapidly accelerates







- 2005 Collections data launched to GBIF, Paleoportal et al.
- 2006 Digital conversion of transaction records accelerates
- 2006 Digital imaging of specimens rapidly accelerates
- 2007 Physical move of collections to West Campus "begins"









- 2005 Collections data launched to GBIF, Paleoportal et al.
- 2006 Digital conversion of transaction records accelerates
- 2006 Digital imaging of specimens rapidly accelerates
- 2007 Physical move of collections to West Campus "begins"
- 2008 Yale Office of Digital Assets & Infrastructure (ODAI)









- 2005 Collections data launched to GBIF, Paleoportal et al.
- 2006 Digital conversion of transaction records accelerates
- 2006 Digital imaging of specimens rapidly accelerates
- 2007 Physical move of collections to West Campus "begins"
- 2008 Yale Office of Digital Assets & Infrastructure (ODAI)
- 2009 Peabody partners w/ODAI & other campus units
- 2009 Physical move of collections to West Campus "ends"











- 2005 Collections data launched to GBIF, Paleoportal et al.
- 2006 Digital conversion of transaction records accelerates
- 2006 Digital imaging of specimens rapidly accelerates
- 2007 Physical move of collections to West Campus "begins"
- 2008 Yale Office of Digital Assets & Infrastructure (ODAI)
- 2009 Peabody partners w/ODAI & other campus units
- 2009 Physical move of collections to West Campus "ends"
- 2010 Yale Digital Asset Management (DAM) system live













- 2005 Collections data launched to GBIF, Paleoportal et al.
- 2006 Digital conversion of transaction records accelerates
- 2006 Digital imaging of specimens rapidly accelerates
- 2007 Physical move of collections to West Campus "begins"
- 2008 Yale Office of Digital Assets & Infrastructure (ODAI)
- 2009 Peabody partners w/ODAI & other campus units
- 2009 Physical move of collections to West Campus "ends"
- 2010 Yale Digital Asset Management (DAM) system live
- 2011 Yale Cross Collections Discovery (CCD) portal live
- 2011 KE EMu interoperating with Yale DAM & CCD











#### Digital History & Timeline



- 2005 Collections data launched to GBIF, Paleoportal et al.
- 2006 Digital conversion of transaction records accelerates
- 2006 Digital imaging of specimens rapidly accelerates
- 2007 Physical move of collections to West Campus "begins"
- 2008 Yale Office of Digital Assets & Infrastructure (ODAI)
- 2009 Peabody partners w/ODAI & other campus units
- 2009 Physical move of collections to West Campus "ends"
- 2010 Yale Digital Asset Management (DAM) system live
- 2011 Yale Cross Collections Discovery (CCD) portal live
- 2011 KE EMu interoperating with Yale DAM & CCD
- 2012 Projects to database & digitize archives commence













#### **Peabody Collections**

Current Digital Snapshot

•	Anthropology	325,000	Lot
•	Botany	400,000	Individual
•	Entomology	450,000	Lot / Individual
•	Invertebrate Paleontology	350,000	Lot
•	Invertebrate Zoology	350,000	Lot
•	Mineralogy	35,000	Individual
•	Paleobotany	150,000	Individual
•	Scientific Instruments	5,000	Individual
•	Vertebrate Paleontology	125,000	Individual
•	Vertebrate Zoology	185,000	Lot / Individual

14 million specimens yields 2.4 million items to catalog electronically



### **Peabody Collections**

Current Digital Snapshot

0	>	70%
0	>	50%
	<	50%

•	Anthropology	325,000	Lot	
•	Botany	400,000	Individual	
•	Entomology	450,000	Lot / Individual	
•	Invertebrate Paleontology	350,000	Lot	
•	Invertebrate Zoology	350,000	Lot	
•	Mineralogy	35,000	Individual	
•	Paleobotany	150,000	Individual	
•	Scientific Instruments	5,000	Individual	
•	Vertebrate Paleontology	125,000	Individual	
•	Vertebrate Zoology	185,000	Lot / Individual	

Items with an electronic record available (25 years' effort): ~55 %



#### Peabody Collections

Current Digital Snapshot

> 70%> 50%< 50%</li>

- Anthropology
- Botany
- Entomology
- Invertebrate Pale
- Invertebrate Zoo
- Mineralogy
- Paleobotany
- Scientific Instrui
- Vertebrate Paleo
- Vertebrate Zoole



230,000 digital assets

mostly JPG & TIF, scattering of other mime types

ot
Idividual
ot / Individual
ot
ot
ot
Idividual
Idividual
Idividual
Idividual
Idividual
Idividual
Idividual

~55 %

Items with an electronic record available (25 years' effort): ~55 %

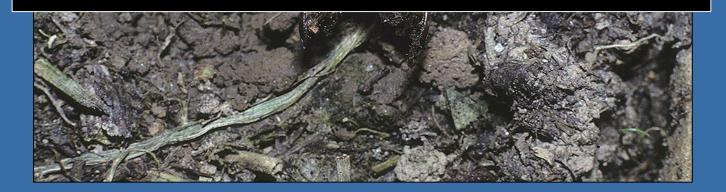




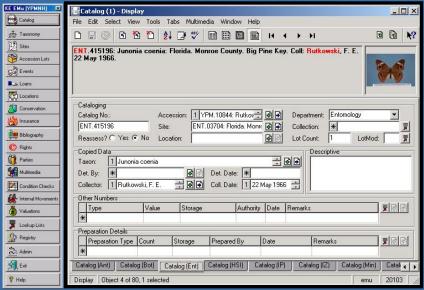


2007

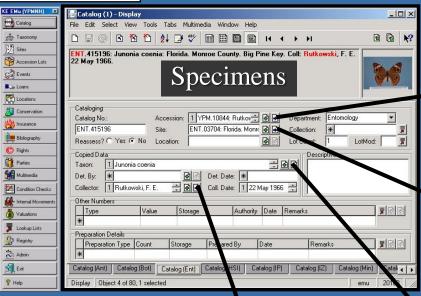
### What to do first?





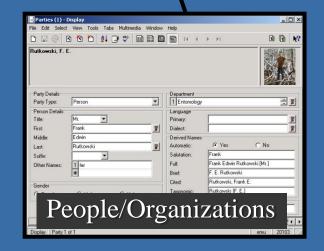


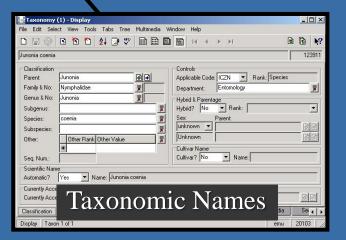




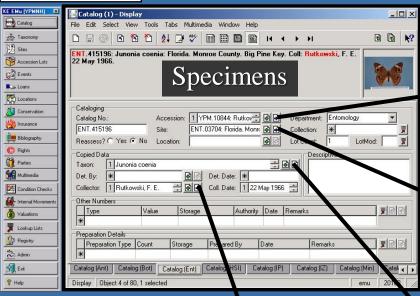


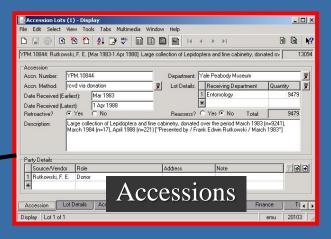


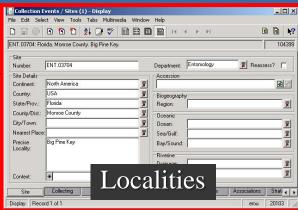


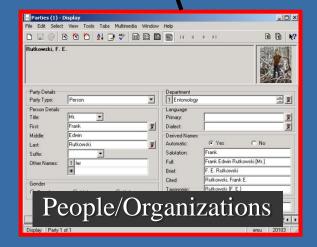


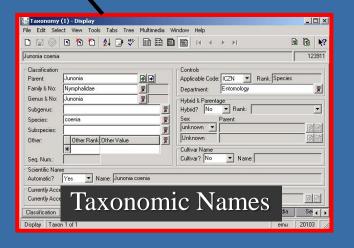




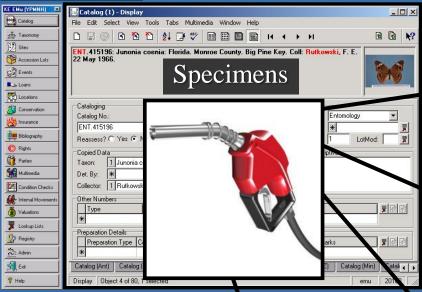






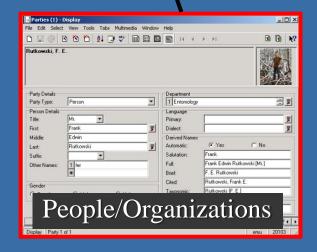


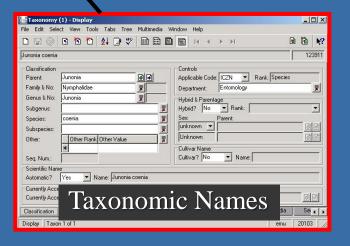










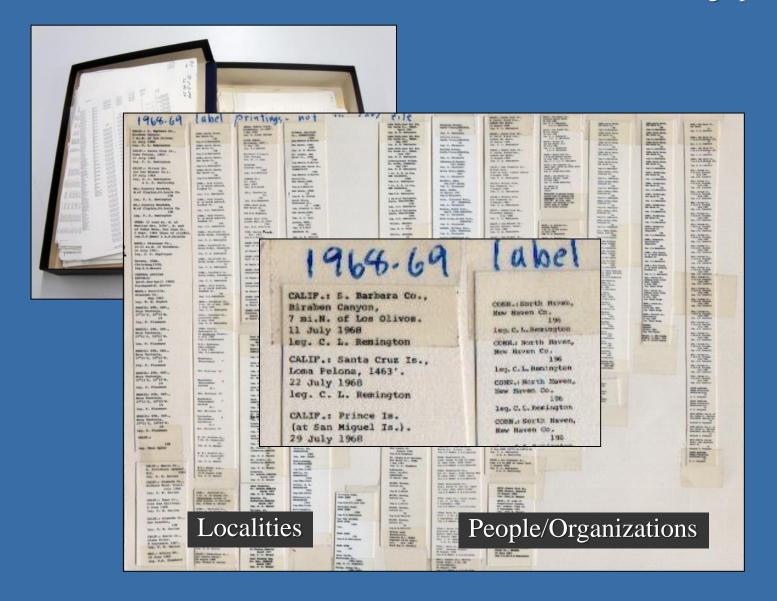






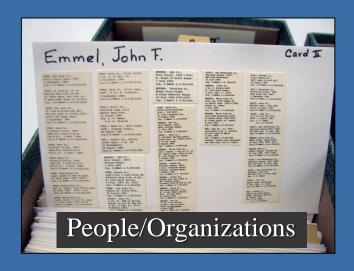














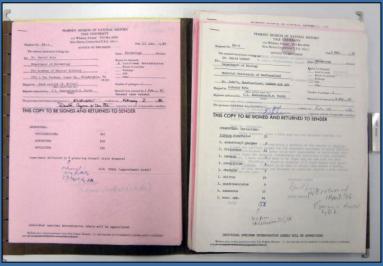


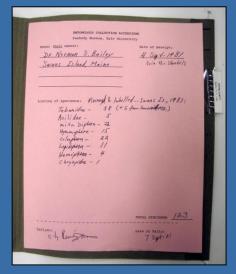










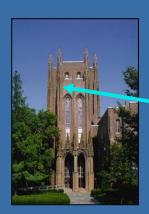


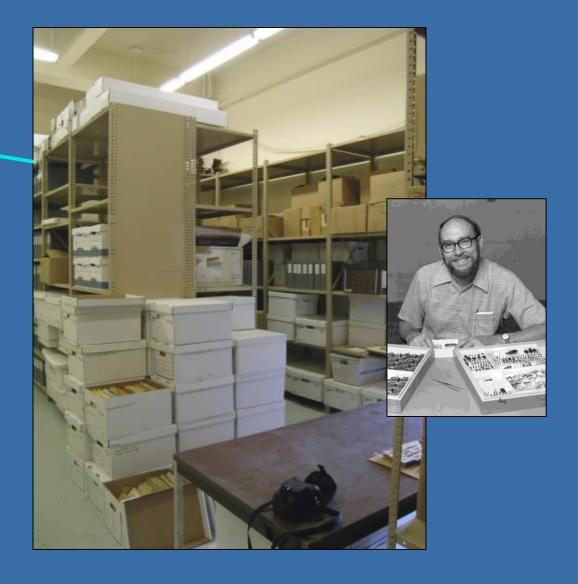










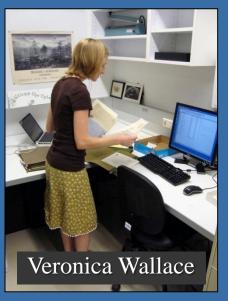




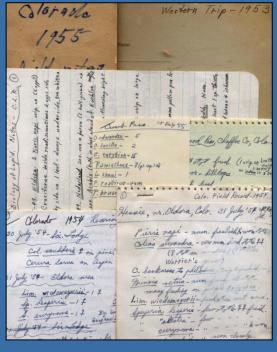




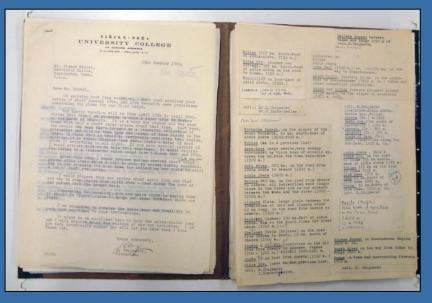


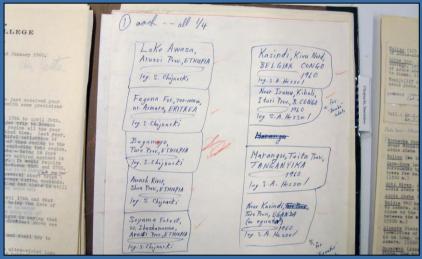


Remington Personal Archives

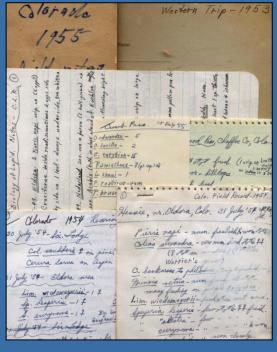




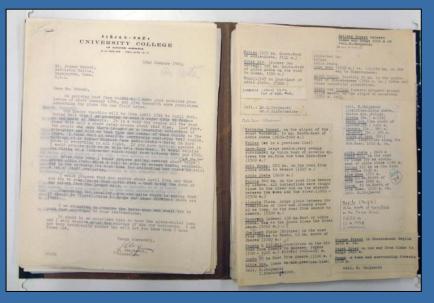


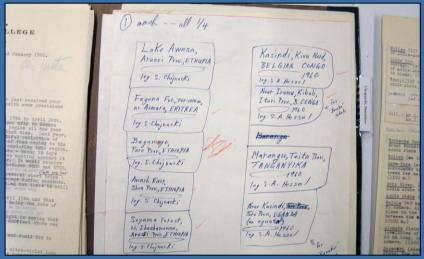






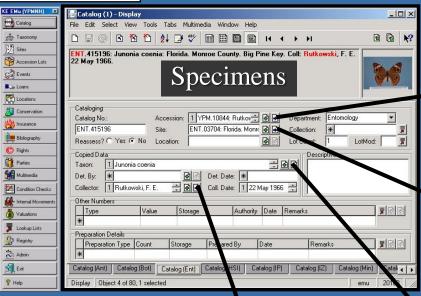






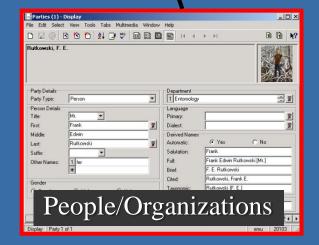


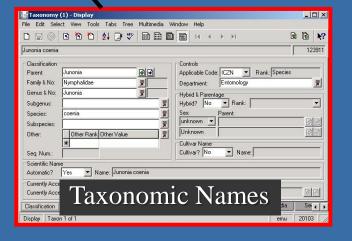






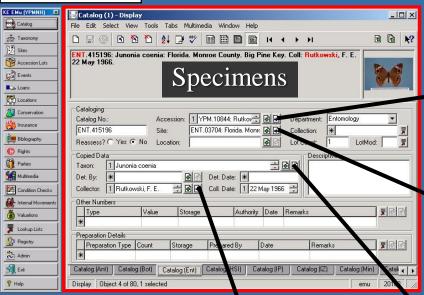






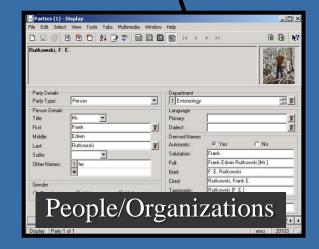


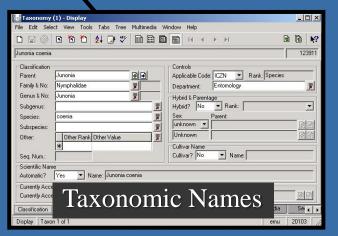




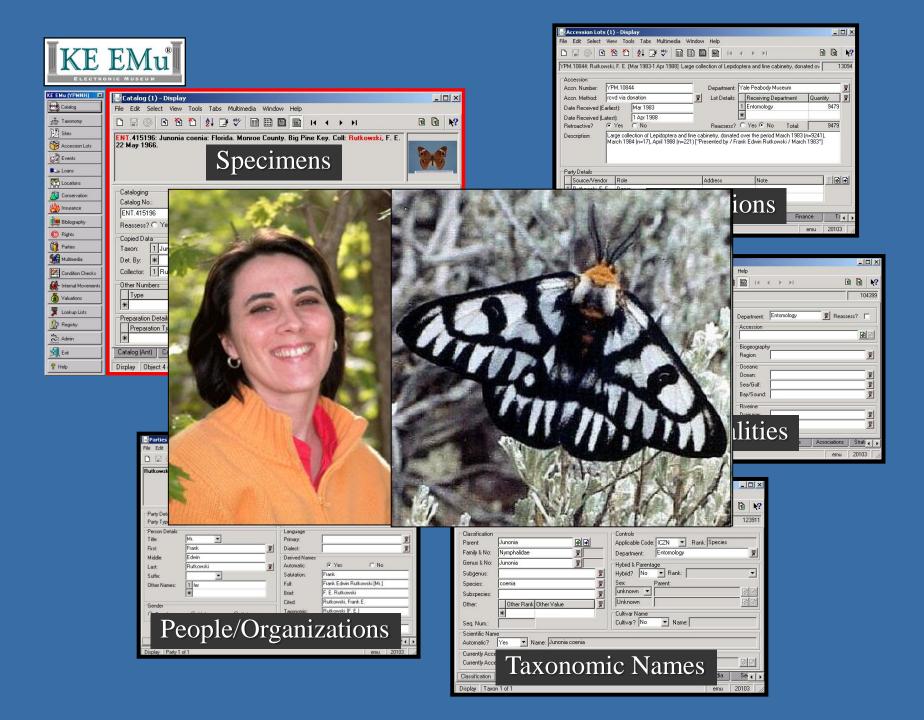














UNCURATED

Sort to Family and Subfamily

Sort to Genus and Species



Sort within Species to Optimize Databasing



Database



Database



Sort to Genus and Species





Database



Sort to Family and Subfamily

Sort to Genus and Species





#### **Student-Powered Paracurating**





Sort to Family and Subfamily

Sort to Genus and Species





FIRST PASS: sort within species to state/province, or "digestible" geographical unit







FIRST PASS: sort within species to state/province, or "digestible" geographical unit



SECOND PASS: sort into "taxon lots" that reflect more fine-grained label metadata



SECOND PASS: sort into "taxon lots" that reflect more fine-grained label metadata



"Read" species in the drawer in columns from left to right, tags demark new species



"Read" species in the drawer in columns from left to right, tags demark new species



Ready for EMu data entry, workroom cleaned, drawers in moveable carts or storage



Ready for EMu data entry, workroom cleaned, drawers in moveable carts or storage

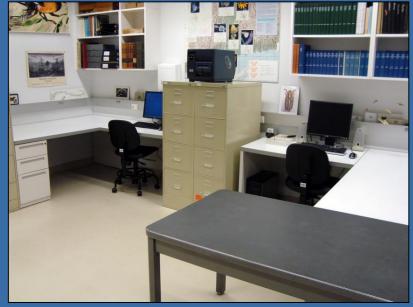


Ready for EMu data entry, workroom cleaned, drawers in moveable carts or storage

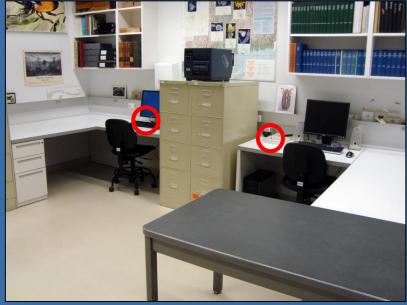


Ready for EMu data entry, workroom cleaned, drawers in moveable carts or storage

















Some aspects of data entry



Some aspects of data entry

29 Aug 2012

#### Cataloging Entomology Specimens in EMu

This is a brief instruction manual for cataloging specimens in KE EMu, the Peabody's collection management system. Most times you will be working from drawers that have been prepared for data entry. If you have any questions, any information seems to be missing, or something just seems wrong, ask Larry. The collections are far from perfect, and deciphering labels is often an intriguing art form, so there will be many questions to ask!

#### Organization of Cornell Drawers

Each Cornell drawer that is ready for data entry will have one or more butterfly or moth species in it. Everything is arranged sequentially in a drawer, usually in four columns of pinning trays of varying sizes. You "read" the drawer's contents by going down each column one by one, starting at the top left of the drawer and ending at the bottom right.

In most cases there will be a handwritten tag with a species name in the top left pinning tray. Depending on whether the drawer contains one or more species, there will be other handwritten tags in the trays. This indicates where one species ends and another begins, as you "read" the drawer. Some drawers contain only one species and so may not have handwritten tags inside, in which case the outside drawer label gives the species name.

Below is a drawer in the process of being databased – there are several different species in the first three columns that have been done, and the fourth column is being worked on (note the yellow tag with the species name on it).

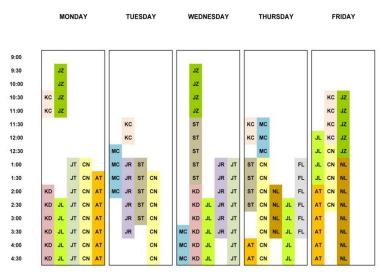


#### About Specimen Lots

Entomology catalogs by "specimen lots" as opposed to by individual specimens. A lot consists of specimens of one species that bear identical label information: all collected at the same place, on the same date(s), by the same person(s), from the same accession, etc. Whenever practical, lots will have been placed into their own pinning tray. In drawers that are ready for databasing, lots are identifiable because their pinning trays have been turned around so that the specimens are upside down. Only one database record is required for an upside down tray. Trays that contain "one-off" specimens are right side up — in such trays, each specimen needs to have its own database record (note that sometimes one-offs have been treated as lots, and are turned upside down; this is usually a judgment call made by Larry to allow some expansion space in the drawer in the future). For each pinning tray, you will need to print and insert a "head label" (see Printing Labels below). The head label for an upside down tray serves as the label for that specimen lot. For a tray of one-offs, you will need to print a head label, but also attach individual pre-printed number labels for as many one-off specimens as there are in the tray.



### **SPRING 2013**



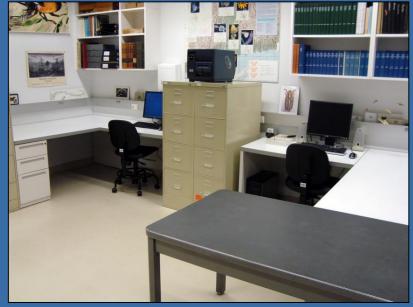


MC=Maria Correa KC=Kathryn Culhane KD=Kristen Dowling NL=Nora Moraga-Lewy

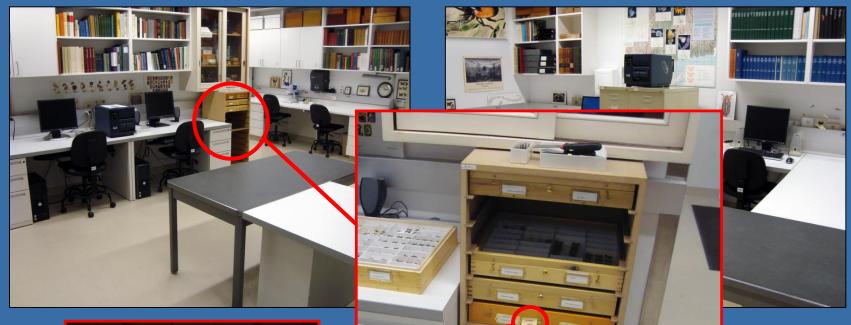
JL=Jessica Lopez NM=Natalie Morningstar CN=Charlotte Newell JR=Jasmine Reid

AT=Alex Terrell ST=Sydney Treuer JT=Jesse Tordoff FL=Florence Loi JZ = Jeff Zirlin













CRITICAL!!

So, how well does this work with students?

# Student-Powered Paracurating 2007-2012 Summary

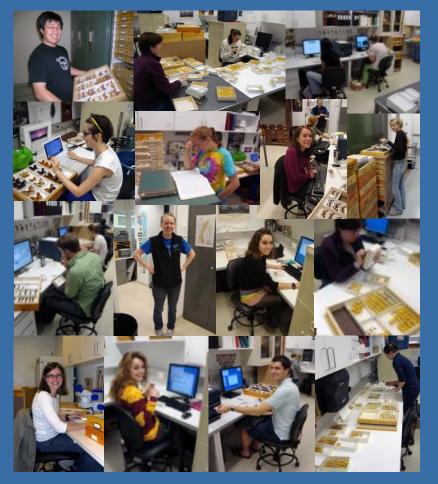
## The Village



57 students

## Student-Powered Paracurating 2007-2012 Summary

## The Village



149,041 lots 273,020 specimens 9,320 hours \$118,471 total

27 specimens / hour \$0.43 / specimen

paracurating & databasing

57 students

## Student-Powered Paracurating 2007-2012 Summary

## The Village



57 students (Provost Office split)

149,041 lots 273,020 specimens 9,320 hours \$118,471 total

27 specimens / hour \$0.43 / specimen

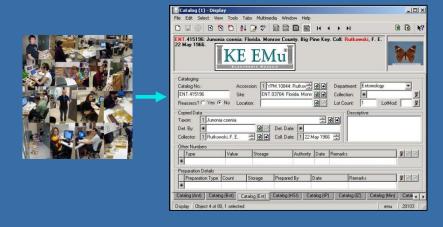
paracurating & databasing

\$59,603 total \$0.22 / specimen





Nicole Palffy-Muhoray Grad. Student, Linguistics

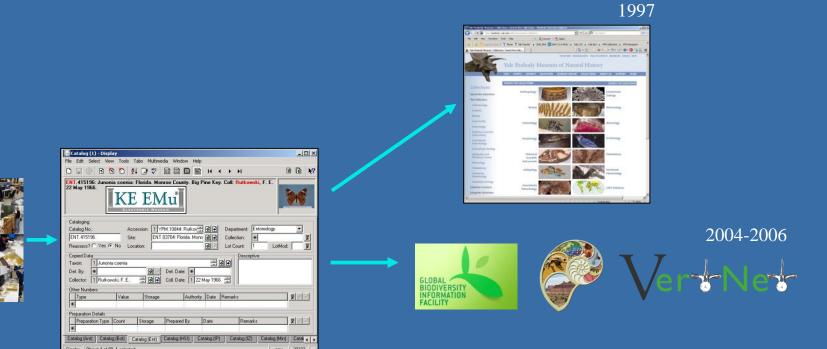


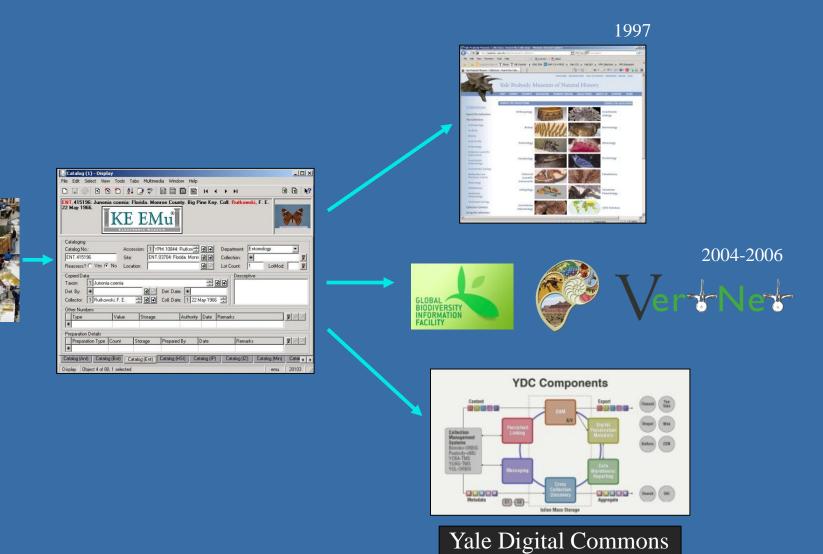
1997



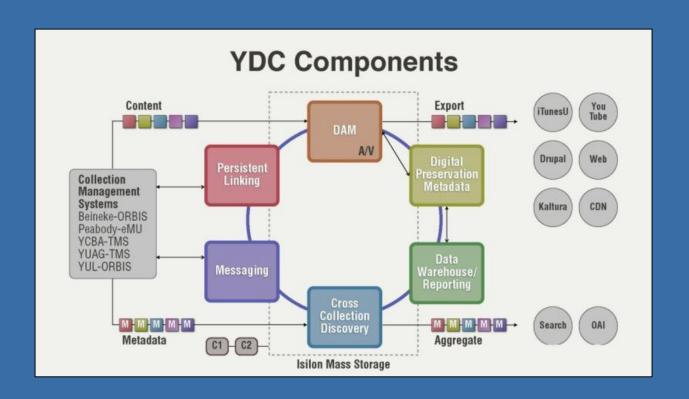


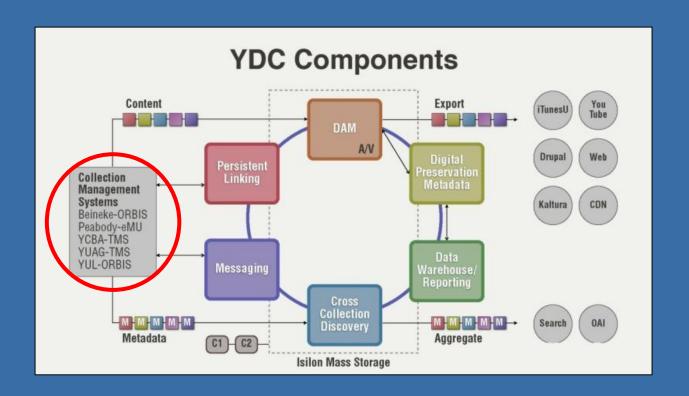
cutulog	(1) - Displ	ay		-						_	□ ×
le Edit	Select View	w Tools Ta	bs Multime	edia Windor	v Help						
	B 6	0 0 21	<b>→</b> **			14	<b>+</b> +	H			₩?
NT.4151 22 May 19	96: Junon 966.		EN	1u <sup>*</sup>	y. Big F	ine Ke	y. Coll	Rutkows	ki, F. E.	30	*
Cataloging											
Catalog N	0.:	Acce	ssion: 1 Y	PM.10844: F	utko	<b>3</b>	Dep	artment: E	ntomology	•	
ENT.415	196	Site:	ENT	.03704: Flori	da. Monti	@ €	Colle	ection:	K		7
Reassess	C Yes G	No Loca	ion:				Lot	Count: 1	LotMo	d.	W
Copied Da	da .							Descripti	VR		
Taxon: 1 Junonia coenia											
Det. Bv.	*		P)	Det. Date:	*	-	-				- 1
Collector:	1 Rutkov	rski, F. E.	- 00	Coll. Date:	1 221	May 196	6 +				- 1
Other Nun	bers							12.			
Туре		Value	Value Storage		Authority E		ate Remarks			7 = 18	21
											_
*											
*  Preparatio	n Details								Remarks		
Preparatio	n Details ation Type	Count	Storage	Prepared	Ву	Date		Remark	0	# G	21
Preparatio		Count	Storage	Prepared I	Ву	Date		Remark	.0	<b>F</b> 0	2
Preparatio Prepar	ation Type	Count				Date Italog (IF	7 C	Remark	Catalog (Mir		





2010-2013

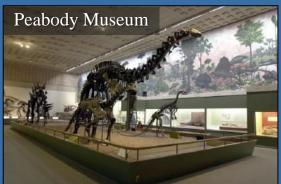


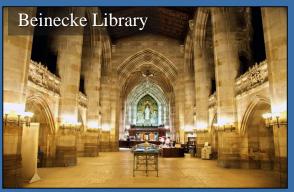












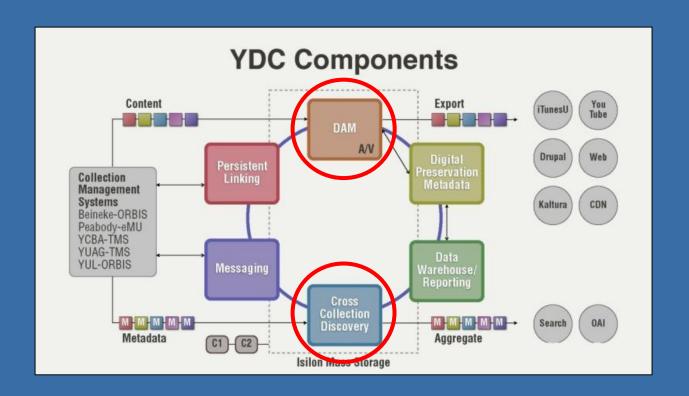








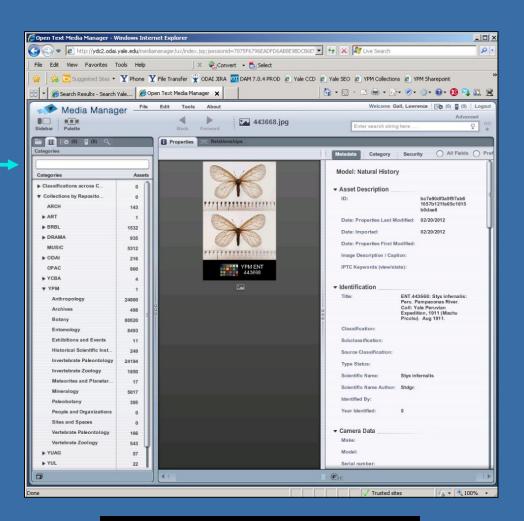








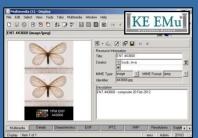




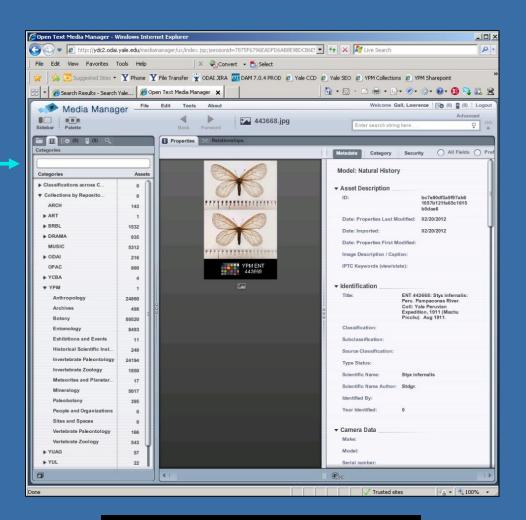
Yale DAM – OpenText Media Manager







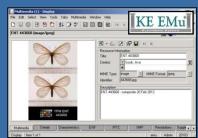


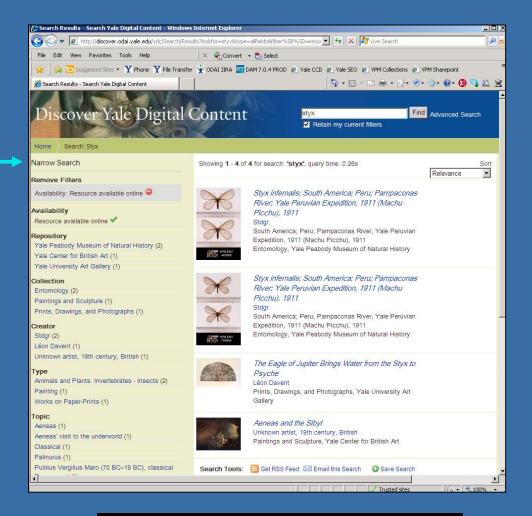


Yale DAM – OpenText Media Manager





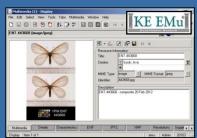




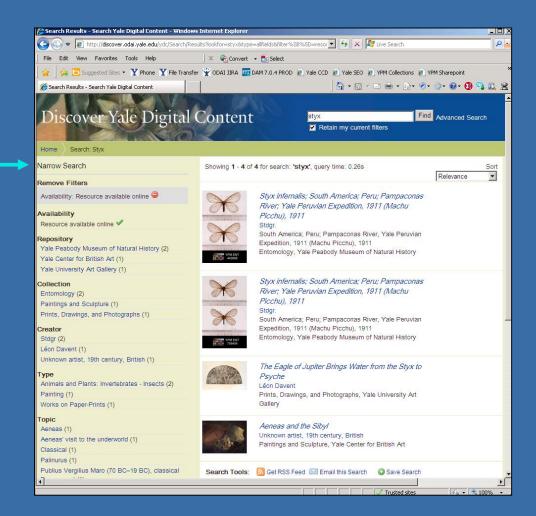
Cross Collections Discovery – VuFind & Solr



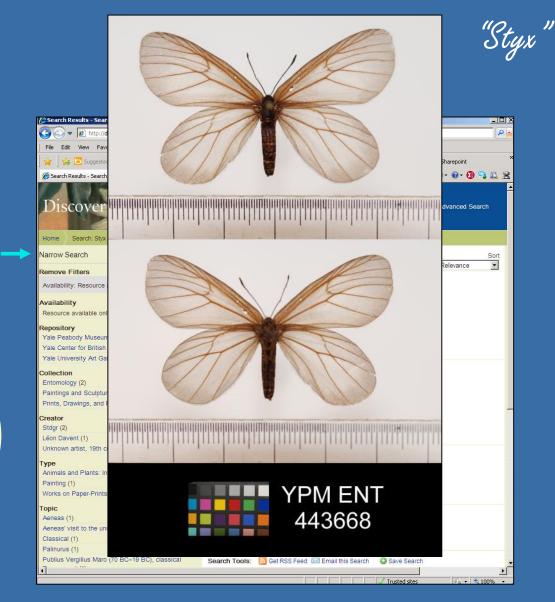




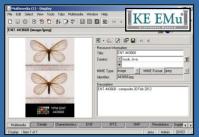




Cross Collections Discovery – VuFind & Solr









Cross Collections Discovery – VuFind & Solr

## The Village



## The Village



First Village Alumna Gilene Young, Yale '07

