

Table of Contents

Vocabulary Program updates Activities overview, LOD, statistics	Online records

Vocab Program Updates

Focusing on TGN, ULAN, and AAT

2012 J. Paul Getty Trust, author: Patricia Harpring .For educational purposes only. Do not distribute.

Recent activities

- Processing contributions and maintaining databases
- Coordinated ongoing Dutch, Spanish, Chinese translation issue resolutions
- Training and coordinating AAT contributors from GCI
- Entering and vetting CONA contributions
- Testing CONA browser and CVA (Contribution Validation Application, formerly called Preprocessor or Processor)
- Research and decisionmaking on LOD issues
- Presented on the Getty vocabularies published as LOD at SAA, AAM, and ALA
- Presented workshop on vocabularies and cataloging in Brazil hosted by Pinocoteca do Estado de São Paolo

Linked Open Data (LOD)

Getty Vocabularies as LOD: A project plan is under development for publishing the four Getty vocabularies as LOD successively, beginning with the AAT. It is anticipated that the data will be published as SKOSextended format under the ODC-BY 1.0 license. The current plan is to also continue providing the data as relational tables and XML. More details, along with information about the time line and status of this project, will be posted on the Vocabulary Program site.

- In 2011 we worked with Marcia Zeng to develop a mapping of AAT to SKOS extended
- In late 2012 we began the 'Getty Vocabularies as Linked Data' project
- We identified our namespace as http://vocab.getty.edu
 We worked with attorney to identify what type of CC license we are intending to use
- We worked with an outside consultant to explore turning the mapping into published linked data
- We are currently working on a prototype for AAT

- Complete the prototype for AAT
- Draft specifications for publishing AAT, TGN, ULAN, and CONA as Linked Open Data
 Explore the possibilities for outsourcing some of the work
- Begin the implementation
- Our current target is to have AAT available as LOD sometime in 2013

Statistics: Current non-candidate records and terms in databases

 Totals of <u>published</u> (non-candidate) records in vocabularies Dec. 2012

records	terms
36,114	244,665
1,241,020	1,799,859
222,851	581,525
1,011	1,887
	36,114 1,241,020 222,851

10(a)3	or more-carra	idates ili 2011
AAT	34,855	222,787
TGN	992,237	1,310,449
ULAN	203.017	540.824

 Much has been accomplished, with small editorial and technical staff, with help from contributors

Total records including candidates

2011 AAT: 51,029 AAT: TGN: 1,432,253 TGN: 1,348,067 ULAN: 249,190 ULAN: 249,240 CONA: 2,653 CONA:

Total terms including candidates

2011 AAT: 267,334 AAT: 247,885 TGN: 2,035,658 TGN: 1,887,748 ULAN: 638,965 ULAN: 636,676 CONA: 4,850 CONA:

Records added or edited since Nov. 2011

AAT: 15,029 TGN: 90,495 **ULAN: 12,093** CONA: 2,579

Nov (18 days) Dec (15 days) Ja ecords 348 294 totions 53 9 records 1,301 937	The said	nys) Feb (19 days) Mar 171 2,063 38 97 175 1,610	ch (20 daj April (19 daj 1,372 21 52 18 1,550 3,024	May (20 day: Jun (1 501 20 218	100 33 3.099	591 1 64	dayi Sep (17 dayi Oc 142 273 55 151 575 263	1 (21 days Nov (19 days 7 578 404 215 149 387 712	7,258
records 1,301 937 cris 24 21	1.0	Nov (18	Dec (15	Jan (19 days)	9	Sep (17	Oct (21	Nov (19 days)	198,583
AAT records	U	348		571		273			2,435 234,115
AT citations		53	9	38		151	215	149	
JLAN records		1,301	937	1,675		263	387	712	
JLAN cits		24	21	1,092		21	20	87	
GN records		48,996	34,004	30,711		5	22	16	
GN citations		3	0	1		10	3	0	
ONA records		9	10	0		1,228	948	811	
CONA citations		2	11	0		6	13	6	
CONA Processor		0	0	0		0	749	199	
otal Edited		50,736	35.286	34,088		1.957	2.935	2.384	

andene of	- antributar	Contribution	total number ress	etal auto morand	acadad ac Nam	NOT
Apr-10	DIE	BAM	Spanish te	rms	30),976
Nov-1) RK	D	Dutch term	ıs	30),868
Jan-12	TEI	L-DAP	Chinese te	rms	3	3,125
Jun-12	GR	ISC	Stub record	ds		60
Jul-12	GR	ISC	Stub record	ds		40
Aug-1	2 GC	I	Conservati	on terms	8	3,441
Oct-12	. VP		Various CO	NA,GCI		754
Nov-1	2 JPC	5M	Stub record	ds		95
ec-12	TEL-	DAP	Chinese tern	าร	received: 5	5,348
	grove Art	artists	10004	10598		
	GRI Library GRI Spec Coll	artists	12860 3057	6322 71		
	ARTstor	artists & repositories	7,071	0		pendi
	GRISC GRISC	stub records	407 406		407 406	pendi pendi

Data Loads as of nber 2012	f <u>contributor</u>	contribution total nu	umber recs total auto-merged Inserted as Ne	W VCS processing note
,	CHIN	Franch tarms	2556	
8	Jun-10	Provenance I	Artists	5,288
0	Sep-10	Fine Arts, IU	Repositories	12,939
1	Nov-10	Grove Art	Artists	16,604
,	Feb-11	GRI Library	Artists	12,860
	7	GRISC	Artists	3,057
	Jan-12	ARTstor	Artists & repositories	7,071
3	Jun-12	GRISC	Stub records	407
	Jul-12	GRISC	Stub records	406
)	Oct-12	VP	Various from CONA	269
	Nov-12	JPGM	Stub records	95
	GRI Spec COII	artists	3057 71	ne ne
	ARTstor	artists & repositories	7,071 0	pending pro
	GRISC GRISC	stub records stub records	407 406	407 pending pro 406 pending pro

Jun-10	NGA/NIMA	Italy	25,985	
Oct-10	NGA/NIMA	Greece	20,461	
		Germany	157,785	cessing
Nov-10	NGA/NIMA	Belgium	20,238	
		Netherlands	17,004	
Feb-11	NGA/NIMA	Mexico	92,880	
		Spain	40,616	
		Chile	32,936	
Sep-11	NGA/NIMA	Taiwan	36,575	
		United Kingdom	25,465	
Jan-12	NGA/NIMA	New Zealand	16,156	
Mar-12	NGA/NIMA	Egypt	24,369	pend
Jun-12	GRISC	Stub records	8	pendi
Oct-12	VP	Various from CONA	82	
Nov-12	JPGM	Stub records	12	
Dec-12	NGA/NIMA	India	48,337	

Records loaded, by contributor

OI					
pr	Apr-12	GRISC	Work records	1,300	
un- ul-	Jun-12	GRISC	Stub records	61	
ul-	Jul-12	GRISC	Stub records	45	
Vov	Oct-12	JPGM	Work records	785	
Apr un- ul- ul- vov		VP	Various for CONA pilot	2,659	
un- ul-	Nov-12	JPGM	Work records		2,672
	Jun-12	GRISC	Stub records	92	
	Jul-12	GRISC	Stub records	102	
	Nov-12	JPGM	Stub records	300	
		The second secon			

- Log of data loaded into CONA and CONA IA
- Column added for CVA/Processor

2012 J. Paul Getty Trust, author: Patricia Harpring .For educational purposes only. Do not distribute

Statistics: Candidates and noncandidates by contributor

ULAN -Total CA	ANDIDATE
ARTstor	6876
AVERY	2437
BHA	139
BVRL-UNM	1
CCA	2
CL-Courtauld	1284
FAL-IU	5
FDA	1
GRISC	2814
GRL	620
GRLPSC	91
Grove Art	3347
JPGM	41
PROV	5778
VP	3294
WCI	19
WCP	26
WL-Courtauld	3766

ULAN candidate and
non-candidate totals

ULAN-NON-CA	NDIDATE
ART500	2
ARTstor	812
AVERY	85840
BHA	35588
BVRL-UNM	1483
CCA	2550
CENSUS	277
CL-Courtauld CLVR	5515
CLVR	2
COBE	2
DHA-	
UCBerkeley	1
FAC-LutherColl	38
FAL-IU	12841
FDA	391
FRICK	5071
GCI	27
GOMRC	2
GRI-DRM	9
GRISC	1104
GRL	14118
GRL-TS	29

GRLPSC

Grove Art	13242
JPGM	2803
LML	1
MAM	1
PROV	13077
Queens	1
RCL	1
SAIC	10
SCA	838
SPB	1
TRAINING	20
USF	1
UTA	2
VP	187025
VP-Intern	31
WCI	18043
WCP	39614
WL-Courtauld	60909

Statistics: Candidates and noncandidates by contributor

ЗНА	3
Bureau AAT	45
CCA	2
CDBP-DIBAM	174
CHIN	12
GCI	15
GRISC	26
CCD	11804
JPGM	34
Suzzallo-UW	7
TRAINING	3
VP	2919
NordImage-V&A	1

- AAT candidate and non-candidate totals
- Note that records are moved to temp.parent rather than being deleted in AAT, e.g., former Guide Terms

30854
30034
31112
2592
3
8243
64
6
353
7
2
59
3125
1
1
36111
1
3
38

AAT -NON-CANDIDATE

© 2012 J. Paul Getty Trust, author: Patricia Harpring .For educational purposes only. Do not distribute.

Statistics: Candidates and noncandidates by contributor

TGN -Total CANDI	DATE
BHA	455
BSR	4
Di Fab	10
FDA	90
GRLPSC	101
VP	193941
TGN -Total NON-C	ANDIDATE
BHA	12327
BPL	3
BSR	19
DHA-UCBerkeley	2
DiFab	5
FDA	872
GRISC	14
GRLPSC	4269
MHS	2
VP	1190582
VP-intern	88
WHS	2

CONA -Total C	CANDIDATE
GRISC	962
JPGM	454
VP	1227
CONA -Total N	NON-CANDIDATE
CAA-AIIS	55
GRISC	319
JPGM	329
Rijksmuseum	99
Te Papa	160
VP	697

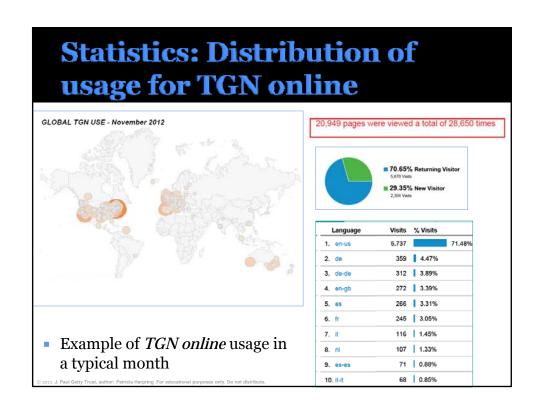
 CONA candidate and non-candidate totals

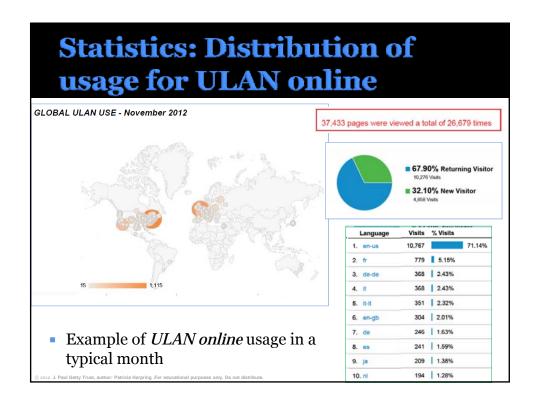
 TGN candidate and non-candidate totals

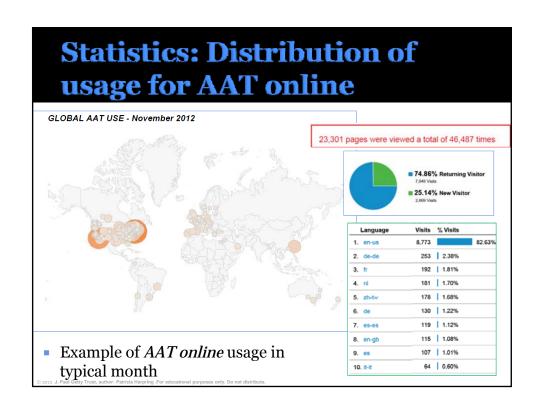
ivity fo	r on	ling	hano		
TATEN TO:	r ou			EATO	
		ше	DIO	M2G	
Browser Reports: Searches and User Activity FY 13	• Con	sistently	high usage	for sear	ching or
Activity F1 13	July-12	August-12	September-12	October-12	November
AAT Page Views	92,698	77,659	82,853	53,419	88,7
TGN Page Views	49,361	41,414	44,805	31,605	47,1
ULAN Page Views	89,439	70,188	72,129	56,479	72,1
Total Page Views:	231,498	189,261	199,787	141,503	207,9
AATUsers	3,525	3,097	3,253	3,965	3,7
TGN Users	3,176	3,280	3,440	3,558	3,1
ULAN Users	5,637	5,594	3,253	6,760	6,4
Total Users:	12,338	11,971	9,946	14,283	13,3
AAT Visits:	10,693	9,528	9,711	11,368	10,6
TGN Visits	8,145	8,150	3,440	8,921	8,0
ULAN Visits	15,427	13,766	14,054	16,703	15,1
Total Visits:	34,265	31,444	27,205	36,992	33,7
AAT Unique Searches	41,090	35,645	37,408	46,080	40,4
AAT Total Searches	47,185	40,454	43,428	53,419	46,4
TGN Unique Searches	27,193	22,837	24,185	27,635	28,6
TGN Total Searches	30,634	25,301	26,965	31,065	25,3
ULAN Unique Searches	55,749	43,675	46,173	52,143	43,7
ULAN Total Searches	60,035	46,863	49,502	56,479	46,6
Total Searches: or aducational nurr	261,886	214,775	227,661	266,821	231.3

		•	le
aantast		Getty Wide-November 2012	
context	No.	Page Name	Page Views
	1	/index.html	144,628
	2	/museum/index.html	80,095
In the context of	3	/visit/index.html	79,145
usage on all pages	4	/visit/hours/index.html	57,174
of the Getty site	5	/vwebv/searchBasic	42,962
	6	/museum/exhibitions/index.html	34,935
	7	/art/index html	34,332
	8	/show_events_list.asp?shcode=15	29,211
	9	/visit/exhibitions/index.html	28,006
	10	/user_login.asp	26,575
	11	/art/exhibitions/pompeii/index.html	24,740
	12	/welcome/index.html	18,959
	13	/my_basket.asp	18,128
	14	/research/tools/vocabularies/ulan/index.html	16,124
	15	/museum/about.html	15,881
	16	/research/tools/vocabularies/aat/index.html	15,295
	17	/vwebv/searchBasic?sk=en_US_getty	14,885
	18	/user_data.asp	13,651
	19	/research/tools/index.html	12,952
	20	/payment asp	12,880
	21	/research/tools/vocabularies/tgn/index.html	12,650
	22	/art/exhibitions/florence/index.html	12,657
	23	/research/tools/vocabularies/index.html	12,071
Trust, author: Patricia Harpring .For educational purposes	only 24	/daily events list.asp	11,560

N ELECTIVE TON		Use in GRI-wide	
context		Top 50 Page Views GRI Wide-November 2012	
COLLCAL	No.	Page Name	Page Views
	1	/research/tools/vocabularies/ulan/index.html	16,124
	2	/research/tools/vocabularies/aat/index.html	15,295
In the context of	3	/research/tools/index.html	12,952
usage on pages	4	/research/tools/vocabularies/tgn/index.html	12,650
of GRI sub-site	5	/research/tools/vocabularies/index.html	12,071
	6	/research/index.html	9,646
	7	/research/tools/bha/index.html	5,244
	8	/museum/research/index.html	4,370
	9	/research/exhibitions_events/exhibitions/dyn_mexico/index.html	3,787
	10	/research/library/index.html	3,430
	11	/research/tools/provenance/index.html	2,594
	12	research/tools/provenance/index.html	1,943
	13	/research/publications/electronic_publications/index.html	1,914
	14	/research/tools/article_databases/index.html	1,730
	15	/research/tools/digital_collections/index.html	1,692
	16	/research/tools/portal/index.html	1,629
	17	/research/tools/photo/index.html	1,628
	18	/research/tools/digital_collections/index.html	1,568
	19	/research/tools/guides_bibliographies/guide	1,355
	20	/research/tools/vocabularies/cona/index.html	1,300
	21	/research/exhibitions_events/index.html	1,237
	22	/research/scholars/years/future.html	1,119
	23	/research/publications/electronic_publications/cdwa/index.html	1,099
12 J. Paul Getty Trust, author: Patricia Harpring .For educationa	24	/research/scholars/index.html	1.023







Number of Licenses

Number of active licenses for the Getty Vocabularies

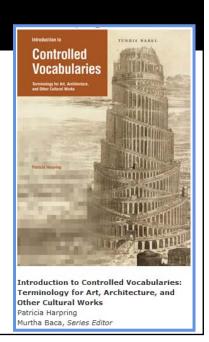
AAT: 86 TGN: 80 ULAN: 60

•Licensing is an issue to be revisited with the adoption of LOD: more later ...

2012 J. Paul Getty Trust, author: Patricia Harpring .For educational purposes only. Do not distribute.

Introduction to Controlled Vocabularies: Revised

- Intro to CV is being updated and published as a revised edition
- Including section on Linked Open Data (LOD)



TGN: Update of recent activity

- Ongoing projects of processing contributions from the National Geospacial Intelligence Agency (NGA, formerly NIMA), GCI archaeological sites, and stubs through CONA
- Outdated administrative hierarchies for various nations in TGN are updated as loads from NGA/NIMA are made

12 J. Paul Getty Trust, author: Patricia Harpring .For educational purposes only. Do not distribute.

TGN: NGA/NIMA loads Example for India: 48,000 new records, 2,000 merged VCS Loader General Info Date Loaded: 12-21-2012 13:32:44 Total Subjects: 48337 Successful: 48337 Failed: 0 Skipped: 0 Inserted: 45765 Merged: 2572 7568055 Bhimg\$01aj Tal\$01av 7568057 Bur\$01al Reef 7568058 Paga Reef 7568058 7568059 Shoal Bank 7568059 New

New

New Coordinates:

Latitude (N, 22?26'4"/22.434444) or longitude (E, 69?55'59"/69.933056) does not

7568060

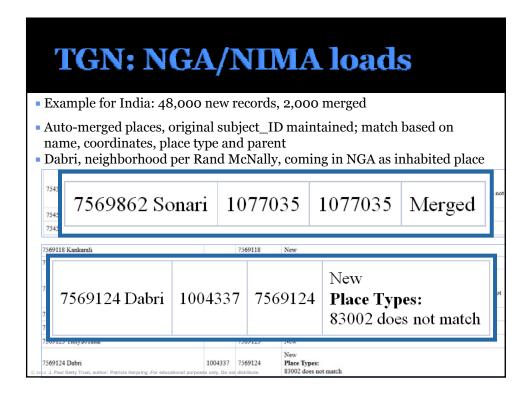
7001508 7568063

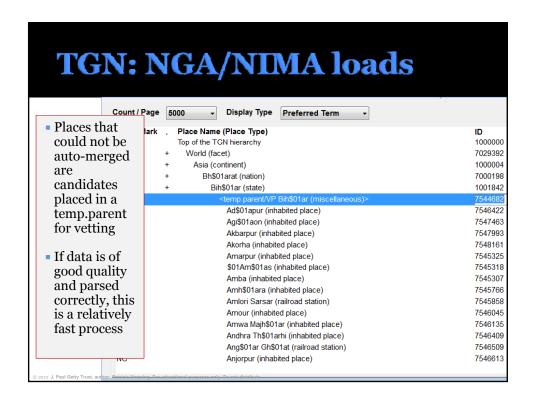
7568060 Pilots Bandar

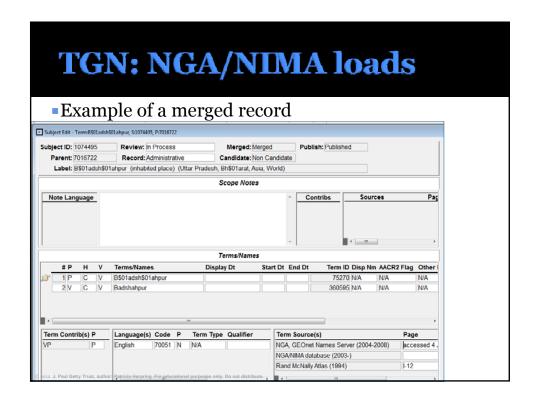
7568061 Malabar Point

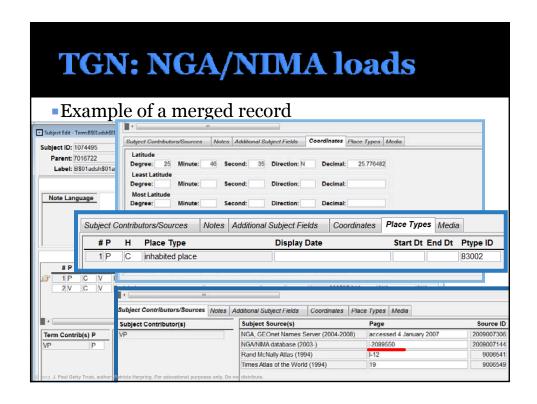
7568062 Sarmat

7568063 Vasai





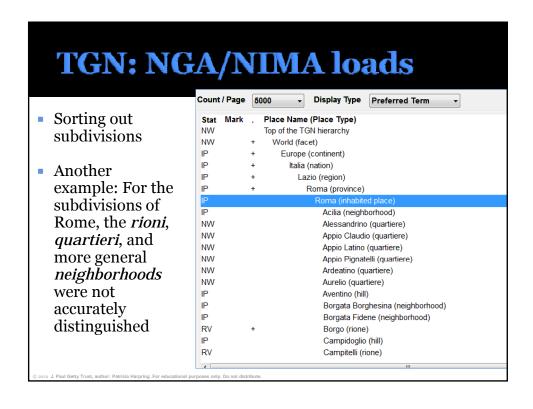




TGN:	NGA/NIMA loads	
 Sorting out subdivisions Do the subdivisions of 		
TGN match the	Mark . Place Name (Place Type)	ID
incoming NGA subdivisions?	Top of the TGN hierarchy + World (facet)	1000000 7029392
0 0110 011 110101101	+ World (facet) + Asia (continent)	1000004
In this case,	+ Bh\$01arat (nation)	7000198
they match	Uttar Pradesh (state)	7016722
fairly well	\$01Ab\$01adgarh (inhabited place)	7568785
For this nation,	\$01Ab\$01adi R\$01ajg\$01ir (inhabited place)	7563273
the second-level	Λbd\$01alpur (inhabited place)	7590332
subdivisions of	Abdull\$01ahnagar (inhabited place)	7559989
these big areas	Abdull\$01ahpur (inhabited place)	7592376
are not	Abdull\$01ahpur (inhabited place)	7592378
included,	Abdull\$01apur (inhabited place)	7563015
causing many	Abdull\$01apur (inhabited place)	7569389
homographs	Abdull\$01apur (inhabited place)	7561878
within a single	Abdull\$01apur Mustaqil (inhabited place)	7564516
level	Abhaipur (inhabited place)	7560618
ICACI	Abhaipur (inhabited place)	7562484
© 2012 J. Paul Getty Trust, author: Patricia Harpr	Abhgaon (inhabited place)	7562423

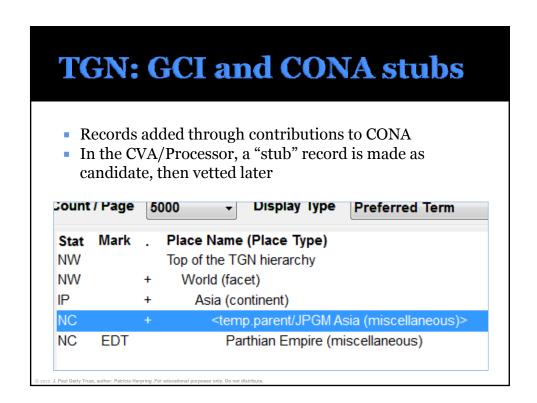
TGN: NGA/NIMA loads

- Even though the match of subdivisions for the India load is fairly good, significant changes have occurred
- Examples of changes to administrative areas that are caught with the NGA load of India
 - state of Mizoram (1001998) changed from a *union territory* to a state since the last time we verified the hierarchy
 - Uttarakhand (7031786) was re-named in 2007 from "Uttaranchal"
 - Puducherry (1002000) was re-named in 2006 from "Pondicherry"
 - Odisha (7016694) was re-named in 2009 from "Orissa"
 - While "Delhi" (7000201) is still the proper short form of the province, its preferred full name changed from the "Union Territory of Delhi" to "National Capital Territory of Delhi" (though it remains a union territory)



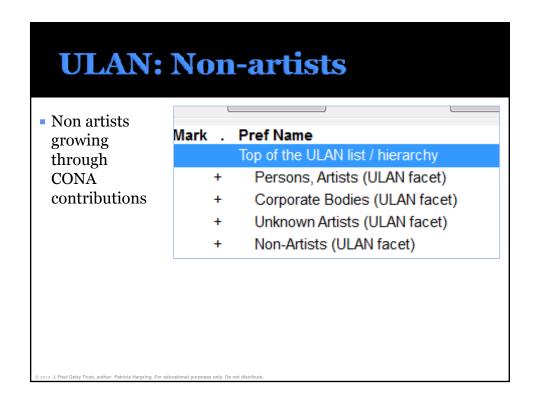


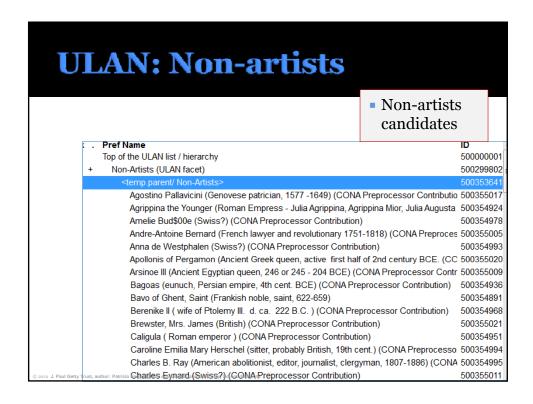
• Examples of records having	EDT EDT EDT EDT EDT EDT	Subject ID 7001372 7522154 1066526 7030405 7001448	Place Name (Place Type) Ab\$01u Ghaush (inhabited place) Ab\$01u Raww\$01ash (ancient site) Anduo (inhabited place) Fuyang (inhabited place)	Parents (Yerushalayim, N (Mu\$05h\$01afaz (Xizang Zizhigu.
records having	EDT EDT EDT EDT EDT EDT	7001372 7522154 1066526 7030405	Ab\$01u Ghaush (inhabited place) Ab\$01u Raww\$01ash (ancient site) Anduo (inhabited place)	(Mu\$05h\$01afaz
records having	EDT EDT EDT	1066526 7030405	Anduo (inhabited place)	•
having	EDT EDT	7030405	. ,	(Xizana Zizhiau
having	EDT		Fuyang (inhabited place)	(Alzany Zizniqu,
9		7001440	,	(Guangxi Zhuang
AATA (CCT)	EDT		Ihn\$01asy\$01a al Mad\$01inah (inhabited place)	(Mu\$05h\$01afaz
AAIAI(TII		7007688	Ko\$13lobrzeg (inhabited place)	(Zachodniopomo
` ' '		7030402	Loulan (former nation/state/empire)	(Taklimakan Sha
48.4		7030404	Maiji Shan (mountain)	(Gansu, Zhonghi
		7032204	Naqa (deserted settlement)	(An-N\$01il, Ash-
Continue		7017002	Neseb\$06ur (inhabited place)	(Burgas, Bulgaria
		7015435	Nin (inhabited place)	(Zadarska \$07Zu
		7030414	Pacatnamu (ancient site)	(La Libertad, Per
■ Added		1106022	Pend\$00eli \$00Oros (mountain range)	(Attik\$00i, Sterea
		7006724	Priverno (inhabited place)	(Latina, Lazio, Ita
names or		1082341	Puy\$06o (inhabited place)	(Ch'ungch'\$06or
		7029596	Quinua (inhabited place)	(Ayacucho, Per\$
continuated		1099878	Qusayr \$67Amra (deserted settlement)	(Amman, Al-Urdu
nort nogonda		7011751	Rumburk (inhabited place)	(Severo\$07cesk
		1025362	Sarand\$04e (inhabited place)	(Sarand\$04e, Sh
		7006443	Sarand\$04e (province)	(Shqip\$04eria, E
		7029578	Shingen Mound (archaeological site)	(Shinshiro, Aichi-
© 2012 J. Paul Getty Trust, author: Patricia Harpring .For edu	ucational n	nurnoses only. Do not	III III III III III III III III III II	

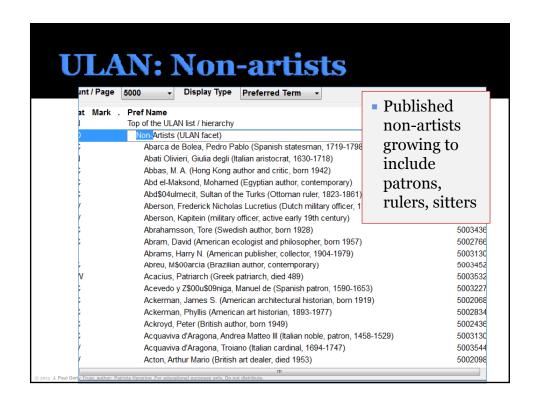


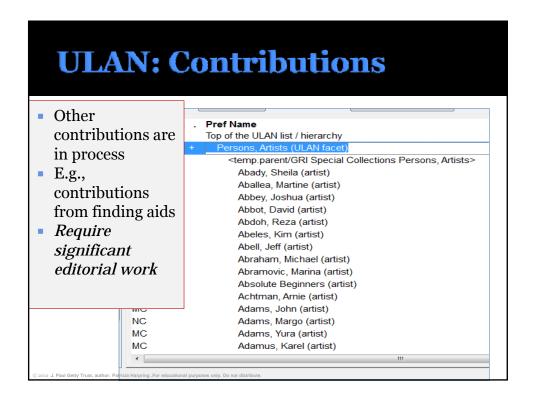
ULAN: Update of recent activity

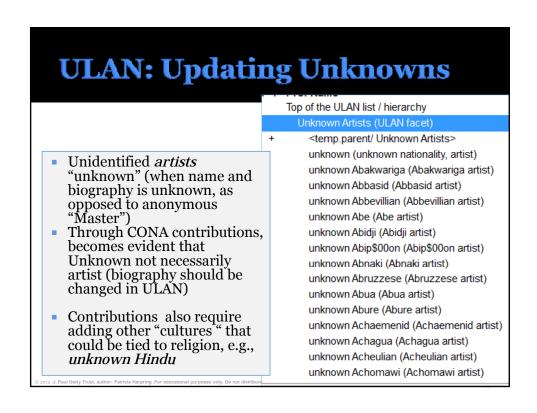
- Additions of non-artists
- Stub records are added and processed via the CONA CVA/Processor
- Ongoing processing of contributions: GRI library, Special Collections, Provenance artists, and Grove Art
- VIAF interaction

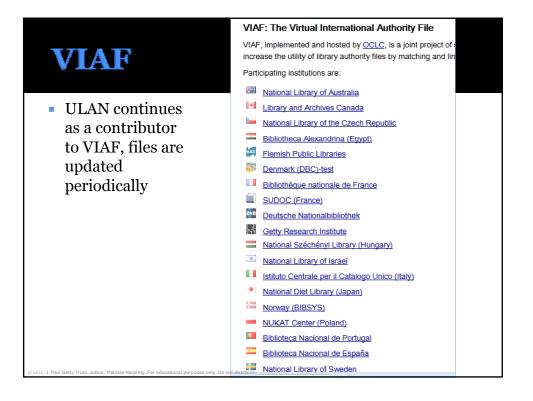


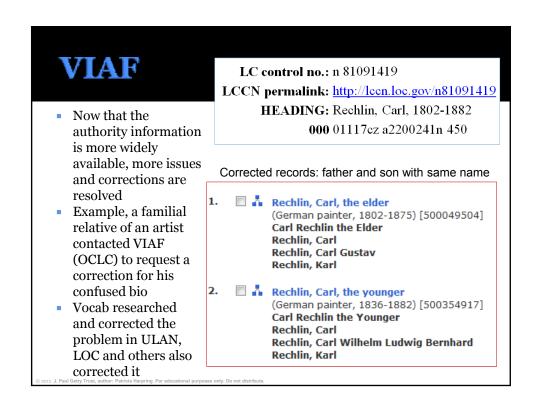












AAT: Update of recent activity

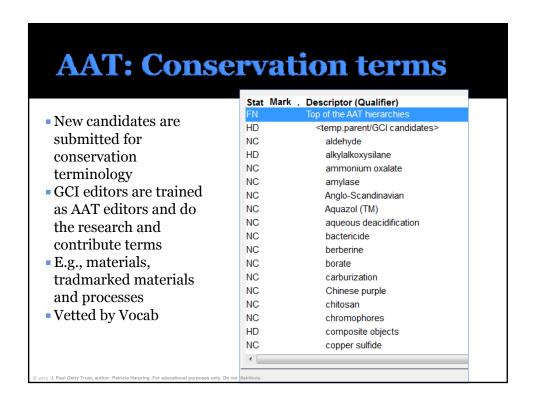
- Processing multilingual loads
- Conservation terms added
- Hierarchies are enlarged as stubs from CONA contributions are processed
- Ongoing cleaning up non-synonyms currently included as UFs in legacy records, adding missing scope notes
- Distinction between Guide Terms and Node Labels is being made more explicit, per suggestions of ISO standards committee

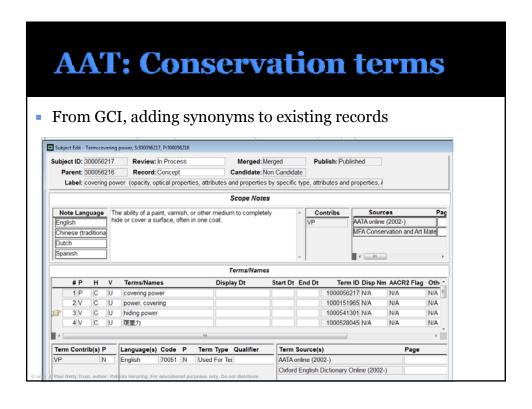
2012 J. Paul Getty Trust, author: Patricia Harpring .For educational purposes only. Do not distribute.

AAT: Multilingual loads

- Multilingual AAT terms:
 - Chinese translation by the TELDAP (Taiwan E-Learning and Digital Archives Program) have so far provided around 3,000 translated records which are online, and around 5,000 recently received but not yet loaded
 - Spanish from Centro de Documentación de Bienes Patrimoniales, Chile; delivered and online
 - Dutch translation from the Rijksbureau voor Kunsthistorische Documentatie was loaded and is online; vetting and edits of legacy Dutch terms from other sources was completed
 - German translation is being undertaken by the Institut für Museumsforschung in Berlin
 - We hope a Portuguese translation will begin soon
 - Full set of 3,000 French terms from CHIN has been fully integrated; occasionally, a plural needs to be added
 - Italian terms from ICCD continue to be vetted

A A TES		•	
AAT: Ca	onservat	ion ter	ms
	Preferred Term	Term	Parent String
 AAT conservation 	35mm	35mm	(small-format, <size: photograp<="" td=""></size:>
• AA1 conservation	35mm cameras	35mm cameras	(<cameras by="" form="" function<="" or="" td=""></cameras>
terminology was	78 rpm records	78 rpm records	(phonograph records, sound re-
harmonized with	abalone shell	abalone shell	(seashell, shell, <shell and="" shell<="" td=""></shell>
	Abbasid	Abbasid	(<islamic dynastic="" styles<="" td="" world=""></islamic>
GCI databases	abbey churches	abbey churches	(<church buildings="" by="" location="" of<="" td=""></church>
New terms being	abbeys (monasteries)	abbeys (monasteries)	(monasteries, religious commu
U	Abelam	Abelam	(<sepik native="" styles="">, Sepik, <</sepik>
added by GCI	abrasion (condition or effect)	abrasion (condition or effect)	*
 Contributions as 	abrasion-resistant material	abrasion-resistant material	(<materials by="" property="">, mater</materials>
	abrasive	abrasive	(<materials by="" function="">, mater</materials>
of December	abrasive cleaning	abrasive cleaning	(cleaning, <subtractive process<="" td=""></subtractive>
2012:	absolute alcohol	absolute alcohol	(ethyl alcohol, alcohol, <oxygen< td=""></oxygen<>
01,		anhydrous alcohol	
Subjects 8,721	absolute dating	absolute dating	(dating, measuring, <analysis a<="" td=""></analysis>
Terms 9,329	absorbency	absorbency	(physical properties, <attributes< td=""></attributes<>
1611115 9,329	absorber	absorber	(sorbent, <materials by="" property<="" td=""></materials>
	absorption (physicochemical proces		
	absorption coefficient	absorption coefficient	(<physical concepts="" sciences=""></physical>
	Abstract (fine arts style)	Abstract (fine arts style)	(<modern arts="" european="" fine="" s<="" td=""></modern>
	Abstract Expressionist	Abstract Expressionist	(<post-1945 and<="" arts="" fine="" styles="" td=""></post-1945>
	abstracting and indexing services	abstracting and indexing serv	,
	abstracts (summaries)	abstracts (summaries)	(summaries, <versions docu<="" of="" td=""></versions>
	acacia (wood)	acacia (wood)	(hardwood, <wood by="" composit<br="">(libraries, <institutions activit<="" by="" td=""></institutions></wood>
2012 J. Paul Getty Trust, author: Patricia Harpring .For educ	academic libraries (institutions)	academic libraries (institution	(libraries, sinstitutions by activit





AAT: Compound terms *flying buttresses* is Contributions from GCI and bound (precoordinated) elsewhere bring up renewed issues re. compound terms Style plus object type is not bound What are the criteria for Baroque + cathedral Baroque + palace precoordination in the thesaurus? Baroque + painting Baroque + concerto When instead should the terms be postcoordinated, combined in search Material plus object type and retrieval? is not bound bronze + sculpture A compound term must express a bronze + bowl single concept or unit of thought bronze + sword capable of being arranged in a genus-species relationship within the AAT hierarchy

AAT: Compound terms

 Animal or plant type plus part of body or product should not be bound (precoordinated)

Stat Mark . Descriptor (Qualifier) FN Top of the AAT hierarchies IΡ Objects Facet IP Components (Hierarchy Name) IΡ components (objects) IΡ <components by specific context>

- goose + feather
- goose + egg goose + beak
- goose + wing oak + leaf
- elm + leaf
- peacock + feather
 If every animal and plant were combined with every possible part or product, would be impractical in the AAT
 - Expanded rules and examples will be developed, discussion with GCI and other contributors
 - Taking into account existing exceptions already in legacy data

ankles (animal or human components) arms (animal or human components) baleen [N] beak (keratinous material) [N] bone (material) [N] chests (animal or human components) egg [N] elbows (animal or human components) <excretions and secretions> [N] faces (animal or human components) feather (material) [N] feet (animal or human components)

AAT Rules for compound terms

A quick test: If either of the components of a compound term expresses different concepts when separate than when they are part of a compound term, bind them together in a compound term. If components of a compound term express the same meaning whether they are individual terms or part of a compound term, do NOT make a compound term.

When to create a compound term

Make a compound term when one or more of the following conditions exist:

- When the meaning of the expression cannot be accurately reconstructed from the constituent independent parts, generally because the modifier or the focus term changes its meaning in the compound expression (e.g., High Gothic).
- When the modifier suggests a resemblance, as in a metaphor, to an unrelated thing or event (e.g., rose windows, barrel organs, spade money, saddlebag houses).
- When the modifier defines a specific shape or type of the focus term (e.g., sleigh beds, basket capitals, window seats, ice houses).
- When the modifier defines the purpose or function of the focus term, and this purpose or function gives the focus term a specific form (e.g., prayer rugs, china cabinets, crib barns, measuring cups).

AAT Rules for compound terms

· Ongoing issue resolutions due to translations, where these rules may not work

When to create a compound term (continued)

- When components of a compound term are too vague to stand on their own (e.g., first drafts, official documents, stone cutters, lost cities).
- When the compound includes relational parts of speech (e.g., chests with drawers, books of hours, towers of silence, chapels of ease).
- When the compound term includes a proper noun or adjective (other than a style term) (e.g., Brewster chairs, Usonian houses, Rangeley boats, Schuetzen rifles).
- When the modifier does not define a subclass of the focus term or when the modifier changes the meaning of the focus term (e.g., paper airplanes, tin soldiers, artificial limbs, Greek fire).
- When the modifier is the governing administrator of the focus term (e.g., cathedral cities, capital cities, army bases, colonial cities).
- When one of the components of a compound term is derived from a guide term (e.g., subway cars, research vessels, emergency currency, clothing stores).
- emergency currency, clothing stores).
 When one or more of the components of the compound term are not AAT descriptors and are not appropriate to add.
- Where there may be a problem in retrieval or vocabulary control, as when splitting the compound term would result in
 confusion in retrieval due to too many false hits (e.g., library science, science library, foreign correspondent, folding
 rules).
- When the distinction between two or more compound terms is confused in the discipline and the distinction would be lost if the terms were not bound as compound terms (Hint: If the scope note distinguishes between the terms, make two compound terms) (e.g., log cabins, log houses, business schools, business colleges, military photography, war photography).
- When the compound term is in a foreign language (e.g., lits à la duchesse, cadavres exquis, p'ai-lous, aes rude).
- When the compound term has a trade mark (e.g., Conte crayons (TM), Plastic Wood (TM), Zonolite Concrete (TM)).
- When the compound term has become so familiar in common use or in the discipline, that it is considered for practical purposes to represent a single concept (e.g., *data processing, art history*).

2012 J. Paul Getty Trust, author: Patricia Harpring .For educational purposes only. Do not distribute.

AAT: Language distribution

December 2012

status in November 2011

Of the total 267,480 terms in AAT

English: terms 140,200

- subjects 38,687 (including American or British) Spanish: terms 53,844
- subjects 31,567 (more, plurals)

Dutch: terms 51,741

subjects 30,972 (fewer, some deleted)

French: terms 6,301

subjects 3,840 (how more terms? next slide)

Italian: terms 13,503

 subjects 12,998 (candidates, unbound compounds, for CONA IA)

British English: terms 3,079

 subjects 1,411 (fewer, changing usage and spelling, e.g., hyphenated words book-shelf; lift vs elevator, or "-or" vs "-our" color vs colour, etc.)

German: terms 1,798

subjects: 1,135 (how more terms?)

Chinese: terms 5,541

subjects 3,135 (including transliterated, etc.)

Portuguese: terms 149

subjects 141

Increasing multilinguality

- TGN and ULAN have many thousands of non-English names, although the languages are not often flagged and thus cannot be counted
- AAT: Base language of AAT is American English, but of the over 247,885 total terms in AAT
 - Spanish 51,001
 - Dutch 51,814
 - French 6,253
 - Italian 13,454
 - British English 3,107
 - German 1,760
 - Chinese 198
 - · Portuguese 135 etc.

(these are total records, including candidates not published; one term may be used in multiple languages)

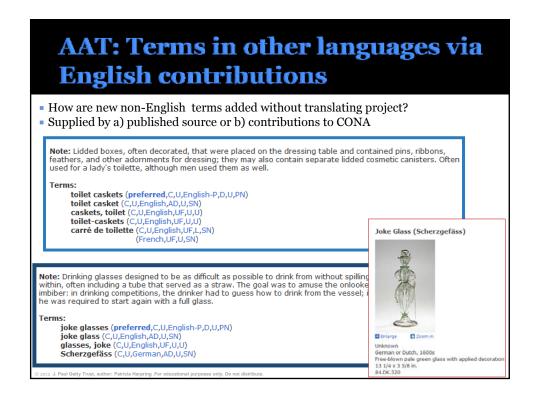
AAT: Terms in other languages via English contributions

- How are new non-English terms added without translating project?
- Supplied by a) published source or b) contributions to CONA
- Legacy data: One primary source of non-English European language terms was IFLA Glossary for Art Librarians, compiled by the International Federation of Library Associations and Institutions, contributed to AAT in 1984 English, Dutch, French, German,

Italian, Spanish,

and Swedish

Record Type: conce artists' proofs (preferred, C, U, English-P, D, U, PN) artist's proof (C,U,English,AD,U,SN) Verbal Communication proofs, artists' (C,U,English,UF,U,N) AP (C,U,English,UF,U,N) ne as the edition, for the use old at inflated prices. Artists' proofs occasionally approaches afdrukken voor de kunstenaar (<u>C,U,Dutch,D,U,U</u>) afdruk voor de kunstenaar (<u>C,U,Dutch,AD,U,U</u>) épreuves d'artiste (C,U,French,D,U,N) (Dutch, UF, U, U) épreuve d'artiste (C,U,Dutch,UF,U,U) (French, AD, U, SN) Künstlerabzug (C,U,German,AD,U,SN) Probeabzug des Künstlers (C,U,German,UF,U,SN) prova d'artista (<u>C,U,Italian,AD,U,SN</u>) pruebas de artista (C,U,Spanish-P,D,U,PN) prueba de artista (C,U,Spanish,AD,U,SN) konstnärsprovtryck (C,U,Swedish,AD,U,SN) provavdrag (C,U,Swedish,UF,U,SN)



AAT: Language distribution

70001 - undetermined 70002 - local language 70048 - Anglo-Saxon 70049 - Old English 70051 - English 70052 - American English 70053 - British English 70055 - Sots Gaelic 70061 - pidgin English 70062 - Gaelic, indigenous 70541 - Polish
 70543 - Portuguese
 70591 - Russian (transliterated)
 70615 - Shuadit
 70643 - Shanish
 70643 - Catalan
 71031 - African language
 71032 - Asian language
 71052 - Asia
 71104 - Arabic (transliterated)
 71171 - Arabic, Tunisian 73171 - Hebrew (transliterated)
 73201 - Hindi (transliterated)
 73202 - Hindiustani
 73370 - Japanese
 73370 - Japanese
 73573 - Khasian
 73573 - Khmer
 73671 - Korean (transliterated)
 73674 - Korean (transliterated)
 73761 - Lakota
 73781 - Lao 111 languages are used by at least one term, among 70062 - Gaelic, indigenous Manx 71171 - Arabic, Tunisian
 71172 - Egyptian
 71173 - Ancient Egyptian (transliterated) 73781 - Lao 73941 - Malayic language several 70101 - German 70102 - High German 73944 - Malay 74291 - Native American languas thousand 70103 - Low German 71181 - Aramaic 71251 - Austronesian language possible 70104 - Standard German 74333 - Nepalese 70241 - Czech 70261 - Dutch 72261 - Bantu 74543 - Onondaga 74690 - Pilipino languages 7499. - Diblina 74769. - Pilipino 74761. - Punjabi 74761. - Punjabi 75174. - Amerindian language 7524.1 - Swahili 75276. - Tagalog 75370. - Tamil 75372. - Tamil (transliterated) 7544. - Thai 75574. - Turkish (transliterated) 7562. - Urdu 7562. - Urdu 7562. - Urdu 75702. - Vietnamese 75808. - Wolof 7594. - Yoruba 72281 - Bella Coola 72310 - Berber 72310 - Berber dialect 72312 - Berber dialect 72371 - Blackfoot 70263 - Afrikaans 70271 - French Many have 70271 - French
70279 - French patois
70379 - Terneh patois
7038 - Sardinian
70409 - Ancient Greek
70410 - Greek
70410 - Greek (transliterated)
70412 - Ancient Greek (transliterated)
70414 - Byzantine Greek
70416 - Liturgical Greek
70416 - Liturgical Latin
70431 - Latin
70431 - Latin
70434 - Medieval Latin
70438 - Magyar (Hungarian)
70511 - Norwegian, New (Nynorsk) 70279 - French patois only one or a "2237 - Blackfoot
 "72431 - Burmese
 72490 - Chadie
 "22550 - Chinese
 "22551 - Chinese (traditional)
 "22552 - Chinese (simplified)
 "22581 - Chinese (transliterated)
 "72581 - Chinese (transliterated Wade-Giles)
 "72582 - Chinese (transliterated Winyin)
 "72584 - Chinese (transliterated Hanyu Pinyin)
 "72584 - Chinese (transliterated Pinyin Wood) few terms linked Among sources for the language list is Ethnologue: Languages of the World and Unicode 72930 - Endo 72952 - Ethiopian documentation 72970 - Persian-Farsi
72972 - Persian-Farsi (transliterated)
72975 - Persian, Middle 70514 - Norwegian, New (Nynorsk) 70515 - Finnish 76048 - Sanskrit76049 - Sanskrit (transliterated) 70517 - Swedish 70531 - Polabian 73135 - Hausa

AAT: Guide terms

- Eliminating unneeded guide terms
- Guide terms become concepts, postable terms
- Guide terms vs node labels; formerly AAT made no distinction; now, at least in definition, they are more clearly explained

guide term

A record representing a concept that serves as a place saver to create a level in the hierarchy under which the *AAT* can collocate narrower concepts (e.g., *<planographic printing processes>*). **Guide terms** should not be used for indexing or cataloging. In *AAT* displays, **guide terms** are enclosed in angled brackets. *See also* **node label**.

node label

A label indicating a division of a hierarchy that contains <u>a characteristic of division</u> (e.g., *<photographs by form>*). Node labels do not represent concepts and therefore have no semantic relationships with concepts. In the AAT, their appearance is the same as guide terms, being enclosed in angled brackets. They should not be used for indexing or cataloging.

AAT: Guide terms become concepts

- Examples of guide terms and node labels
- In AAT, will continue to flag them as Guide Terms

Subject ID Descriptor (Qualifier)

300000763 <districts by condition> 300000768 <administrative bodies>

300000796 <religious districts>

300000900 <corner buttresses>

300000955 <piers by form>

300000966 <piers by location or context>

• Some may become Concepts, as happens when users request it, e.g., *corner buttresses*

2012 J. Paul Getty Trust, author: Patricia Harpring .For educational purposes only. Do not distribute

AAT: Too many divisions

100

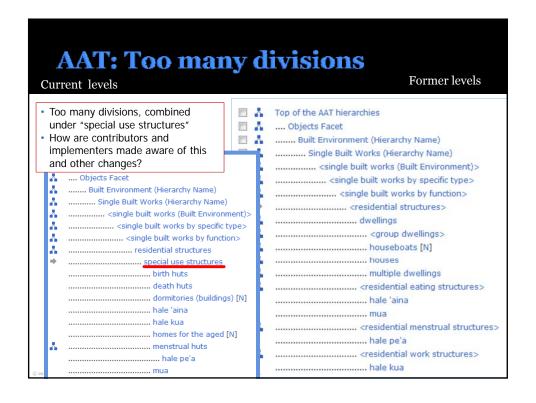
Former levels

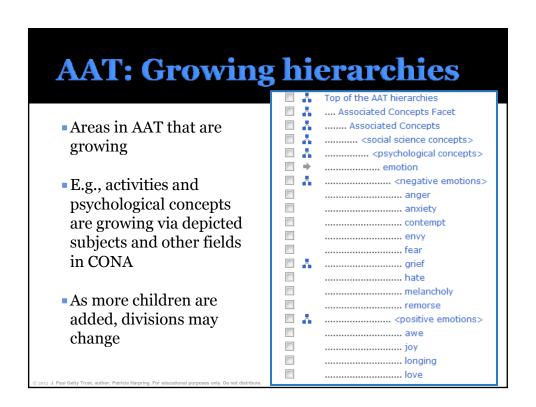
- Example of eliminating guide terms or changing them to concepts, per AAT editorial rules
- A division having only one or two children is not helpful; legacy data issue

Avoid making overly complex divisions that cause unnecessary complexity in the structure; such divisions hinder the ability of end users to access the data through browsing the hierarchies, in addition to making displays difficult to read.

San.		Top of the AAT hierarchies
	٨	Objects Facet
		Built Environment (Hierarchy Name)
		Single Built Works (Hierarchy Name)
		<single (built="" built="" environment)="" works=""></single>
177		<single built="" by="" specific="" type="" works=""></single>
	A	<single built="" by="" function="" works=""></single>
	nþ	<residential structures=""></residential>
[77]		dwellings
(87)		<group dwellings=""></group>
[77]		houseboats [N]
	٨	houses
		multiple dwellings
	Å	<residential eating="" structures=""></residential>
(17)		hale 'aina
		mua
m		cresidential menetrual structures

......<residential work structures>





AAT: Disentangling quasisynonyms or related terms

- Continuing effort to create separate records for nonsynonymous UFs
- Corrected as users point them out
- E.g., "Masses";
 Masses are eucharists, but not all eucharists are
 Masses

aucharists (<religious ceremonies>, ceremonies, ... Events)

Note: A type of sacrament and central act of worship for most Christian denominations that is b Supper when Jesus identified the wine he poured and the bread he broke as his blood and bod generally consists of the consecration of bread and wine by the minister or priest and its distrib worshippers. Aspects of the theology of the eucharist have long been controversial, particularly Reformation. It is now widely agreed that the eucharist commemorates the unique sacrifice of C as to make its effects real. The term derives from the Greek 'eucharistia' meaning 'thanksgiving.

Terms:

ID: 300264326

eucharists (preferred,C,U,English-P,D,U,U) eucharist (C,U,English,DF,U,U) Masses (C,U,English,UF,U,U) Holy Communions (C,U,English,UF,U,U) eucharistic (C,U,English,UF,U,U) eucharistic (C,U,English,UF,U,U) communions (C,U,English,UF,U,U) eucharistievieringen (C,U,Dutch-P,D,U,U) mis (C,U,Dutch,AD,U,U) eucaristias (C,U,Spanish-P,D,U,PN) eucaristia (C,U,Spanish,AD,U,SN) Eucaristia (C,U,Spanish,UF,U,U)

"masses" as separate record should be linked as genus/species, or associative relationship "meaning/usage overlaps"

AAT: Ongoing issue of male/female in separate records

- AAT, as originally designed monolingual, was inconsistent:
 - · actors and actresses in the same record as synonyms
 - but other male and female roles in separate records (duke and duchess)
- Going forward with translations, this becomes a bigger problem, because the distinction between male and female will be different in different languages
- Eventually merge existing English inconsistent male and female records into single records, to be considered synonyms rather than related concepts
- Also better for retrieval if they are in the same records

archivists (preferred,C,U,LC,English-P,D,U,PN)
archivists (C,U,English,AD,U,SN)
archivists' (C,U,English,AD,U,SN)
archivist's (C,U,English,AD,U,N)
archivist's (C,U,English,AD,U,N)
archivist's (C,U,English,AD,U,N)
archivist's (C,U,Chinese (traditional)-P,D,U,U)
archivist's (C,U,Chinese (traditional)-AD,U,U)
ding an bao guan yuan (C,U,Chinese (transliterated Hanyu Pinyin)-P,D,U,U)
dang an bao guan yuan (C,U,Chinese (transliterated Pinyin without tones)-P,D,U,U)
anchivarisesen (C,U,D,Urch AD,U,SN)
archivarisese (C,U,D,Urch AD,U,SN)
archivarisese (C,U,D,Ernah,AD,U,SN)
Archivar (C,U,German,AD,U,SN)
Archivarin (C,U,D,Smaihsh,AD,U,SN)

CONA Update

Contributing, processing, publishing

2012 J. Paul Getty Trust, author: Patricia Harpring .For educational purposes only. Do not distribute.

CONA: Update of recent activity

- The CVA (Contribution Validation Application) was finished and used extensively by the Vocabulary Program
- We tested the CONA online search, released in pilot project but unadvertised due to little data
- Ongoing efforts to bring in contributions
- Refine editorial rules for CONA, examples from real data

CONA: Contributions

How can small staff process so many contributions?

- By carefully mapping each contributor's data to the Vocabularies' data models
 - Sometimes a good fit (museum data), other times requires significant global manipulation by Gregg or a consultant, and manual work by editors where there is no pattern to allow global edits
 - With full data from reliable contributor, auto-merged or added with little intervention and can be quickly vetted
 - With incomplete or inconsistent contributor data, requires more oversight by VP and candidates take longer to vet
- By having a dedicated ITS development team; systems built and maintained to make our editorial work faster and more efficient
- By setting priorities and having quotas: don't spend too much time on one difficult record, but instead complete ten easier records

© 2012 J. Paul Getty Trust, author: Patricia Harpring .For educational purposes only. Do not distribute.

CONA: Contributions

- Among the CONA data sets currently in editorial processing
 - 1,200 records for prints for CONA from Special Collections
 - 3,000 records from JPGM
 - 100 records from the Rijksmuseum
 - 160 records from Te Papa New Zealand
 - Small samples from Palace Museum Taipei, Berlin museums, the Fine Arts Museum in Santiago, and the AIIS (American Institute of Indian Studies) project for Indian architecture
- We will welcome more contributions

CONA: Contribution formats

- CONA XML preferred format, also LIDO or CDWA Lite
- Back and forth between Gregg , Vocabulary Program and the contributor re. the parsing and format
- By necessity, we may accept others and parse them here

```
<?xml version="1.0" encoding="utf-8" ?>
- RijksmuseumXML>
- - decordList>
- - decord_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_steend_
```

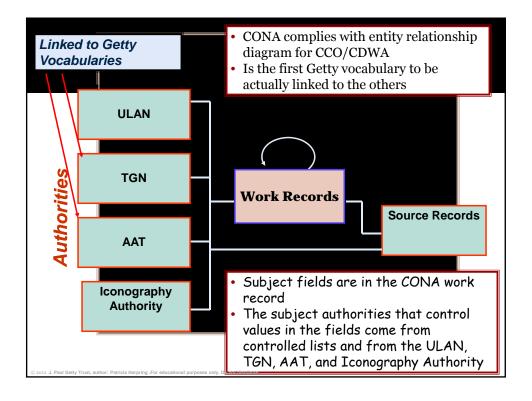
CONA: Parsing contributions

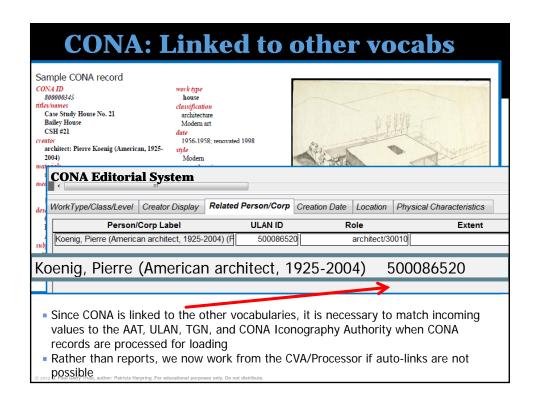
- Every contribution has issues
- If issues in parsing are resolved, the contribution can be vetted and published faster

```
Issue in contributed data
```

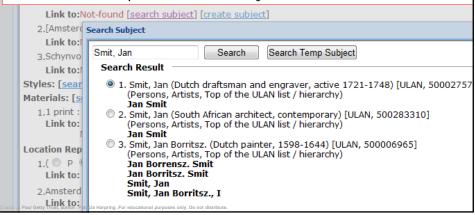
Please parse like this instead

```
<Creation_date> ca. 3000 BCE </Creation_date>
<Start_date> -3100</Start_date>
<End_date>-2900</End_date>
```



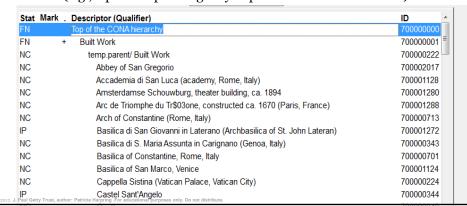


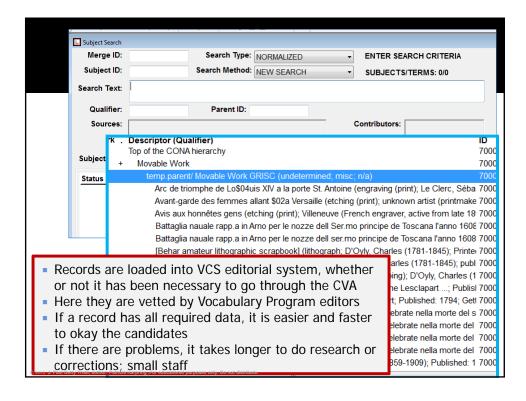
- Unresolved auto-matches to other vocabularies are vetted in the CVA/Processor
- Editor may search other vocabularies in CONA CVA
 - CONA CVA presents editor with choices for linking
- New stub records for AAT, ULAN, TGN, or IA, may be added and linked on the spot, filled in later by editors
- For links with a pattern, editor may write a 'rule' for CVA

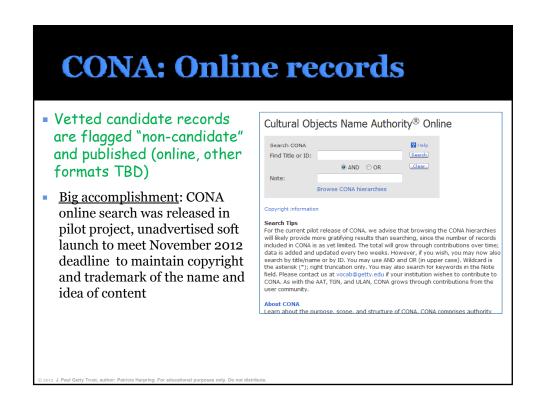


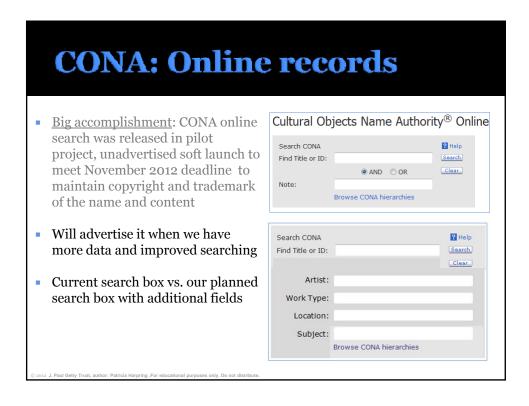
CONA: Stubs

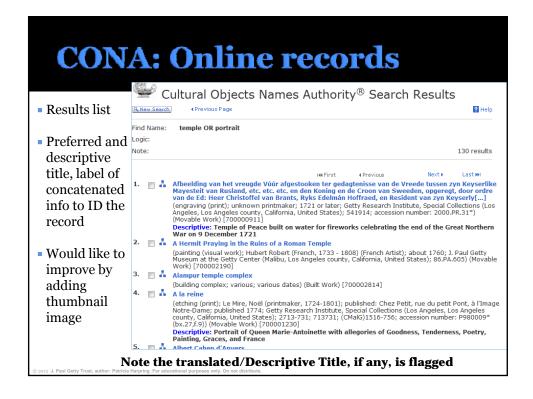
- In CVA, stub records may be made in AAT, TGN, and ULAN, also in CONA itself or the CONA IA
- A stub record is a brief, incomplete record that must be filled in at a later date
- Examples of architecture as depicted subjects in contributed CONA work records (e.g., a print or painting may depict a work of architecture)

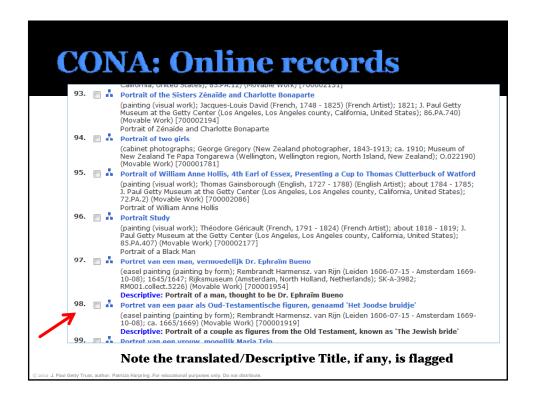














Minimum information

- Catalog Level (item, group, etc.)
- Object/Work Type
- Title or Name
- Creator
- Creation Date
- Measurements
- Materials and Techniques
- Depicted Subject
- Current Location
- Repository number for movable works
- Sources

Default values are available for missing required information. E.g., "unavailable" for measurements.

Minimum information

- Preferred title: The preferred title should be a recent title provided by the owning institution, a concise inscribed title, or a title provided by the artist, if known and if it is sufficiently descriptive
- Exception: If the work is commonly known by a title, use that title as preferred, even if it is not descriptive
- Descriptive Title: Included for all works, if possible. Descriptive Title
 conveys to the end user what the work is or what its subject is about.
 (The subject should also be indexed in Depicted Subject.) If the
 preferred title is not descriptive and in the language of the catalog
 record (English for CONA), a descriptive title should be constructed
- In CONA, the Vocabulary Program will typically not construct a descriptive title but we hope the contributor will supply one as needed
- Typically not an issue with Museum data
- But language = English may be an issue

Merged information

- As with AAT, TGN, and ULAN, data in CONA will be merged when there are multiple contributors for the same record
- When records are merged, one is Dominant, and the other(s) is Recessive
- The defaults and preferences of the Dominant are maintained in the merge
- For a few repeating fields, information from both Dominant and Recessive are included
- In all cases, both contributors and their sources are included
- For movable works, the data from the owning institution will always be Dominant

Creator Display:

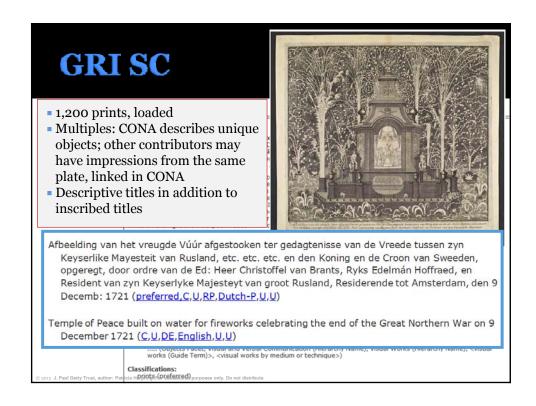
Vincent van Gogh (Dutch, 1853 - 1890) (Dutch Artist) [preferred, JPGM] Vincent van Gogh (Dutch, 1853-1890, active in France) [,VP] Gogh, Vincent van (Dutch painter and draftsman, 1853-1890) [500115588]

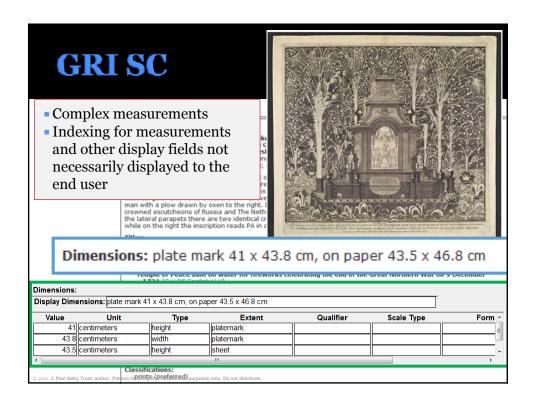
CONA: Contribution examples

- GRISC
- JPGM
- Te Papa
- Rijksmuseum
- AIIS architecture
- Berlin museums, Chilean museums, Taiwanese museums
- The following examples illustrate a few interesting issues and data displays in CONA online

2012 J. Paul Getty Trust, author: Patricia Harpring .For educational purposes only. Do not distribute.

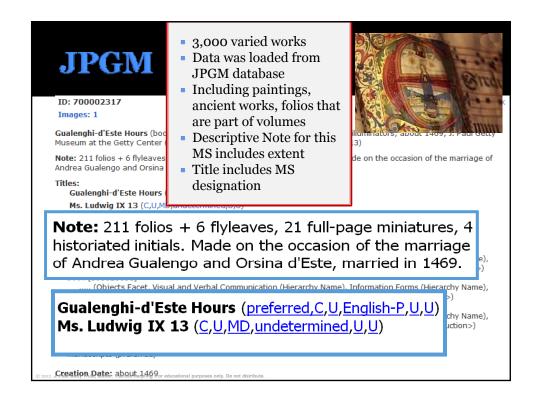




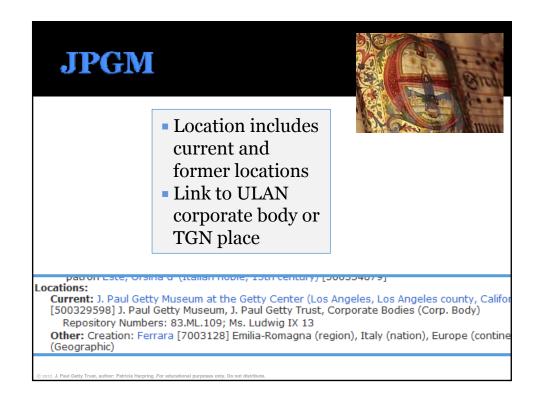


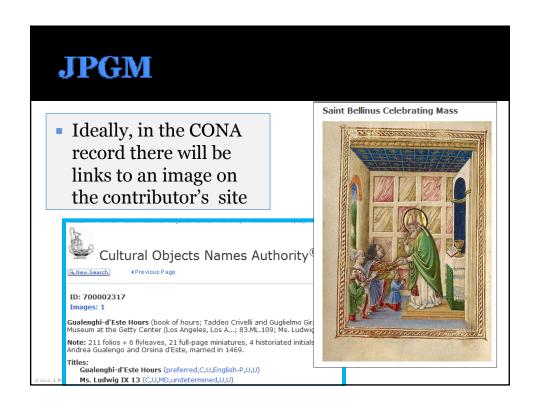


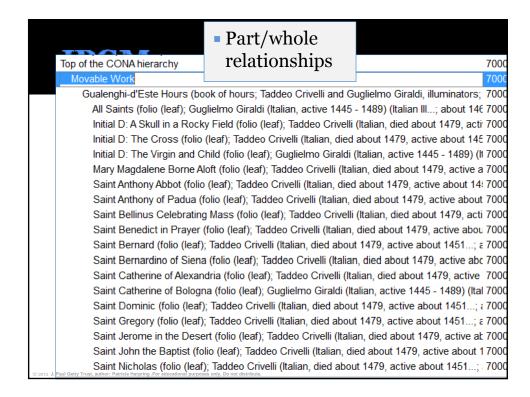


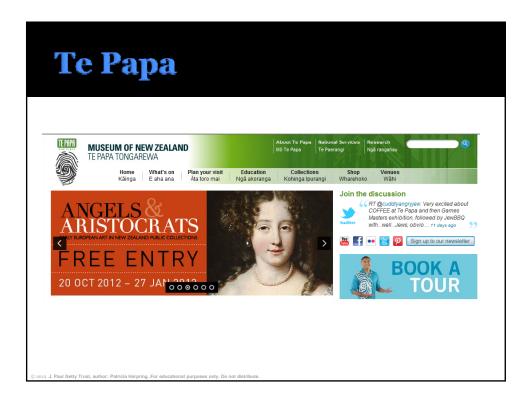


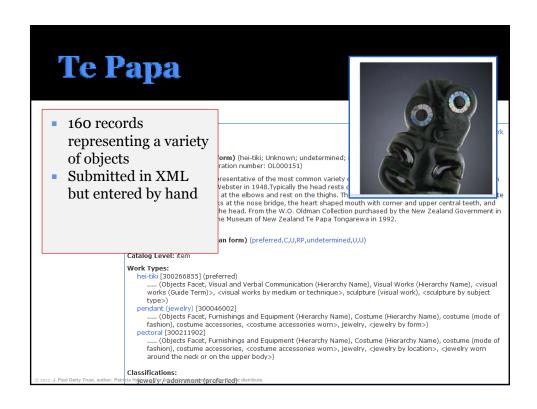


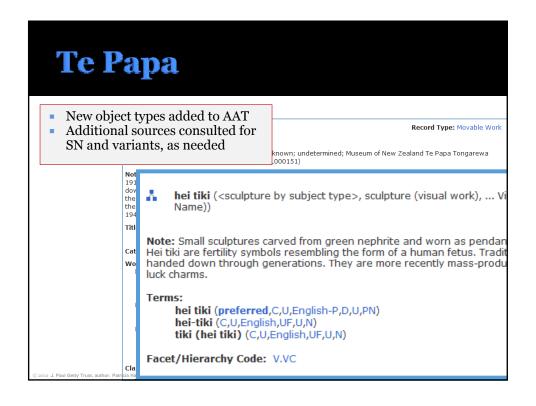


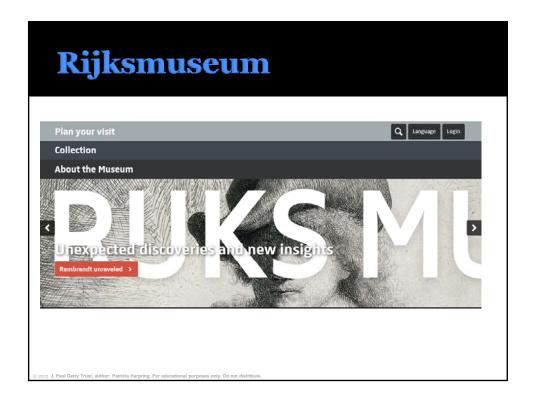


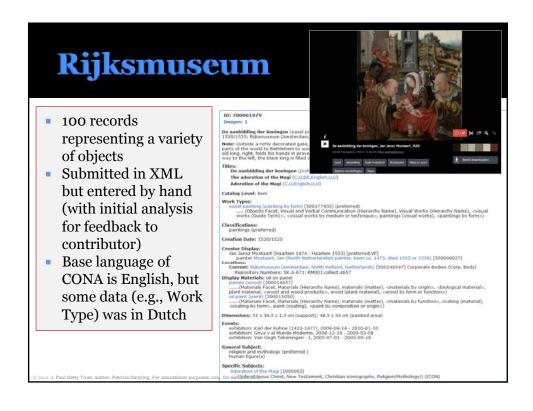


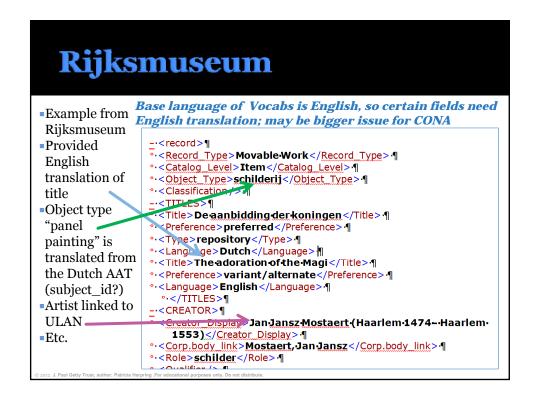






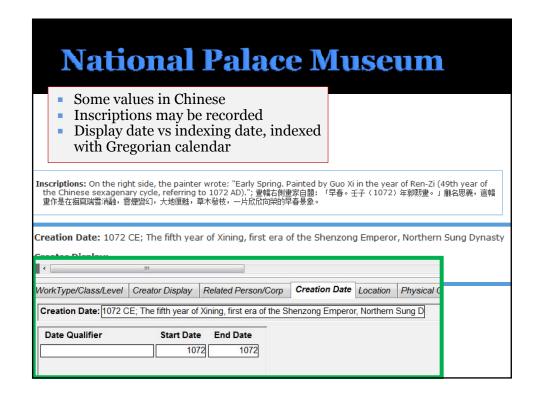




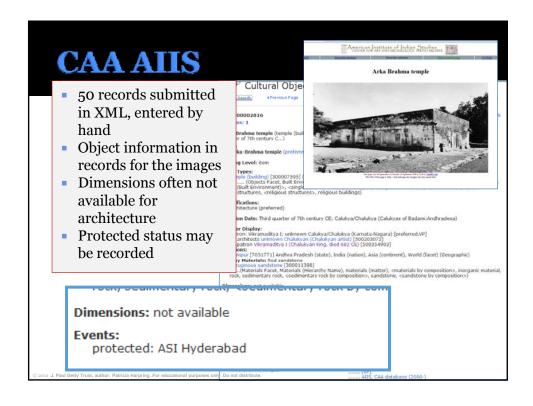


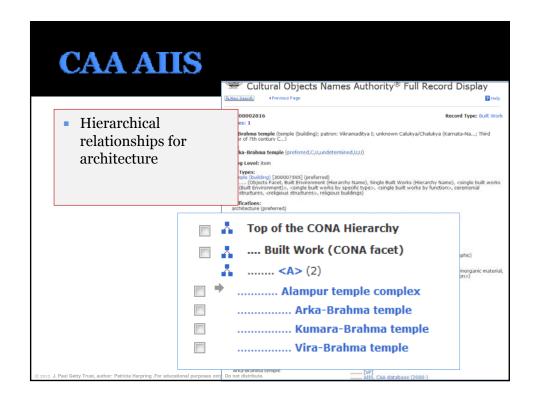


























CONA: Contribution issues

Primary issues in contributed data so far

- Linked data is not in the other vocabulary: solved via CVA
- Missing fields
 - Descriptive titles
 - Depicted subject

2012 J. Paul Getty Trust, author: Patricia Harpring .For educational purposes only. Do not distribut

What is Subject?

- The subject matter of a work of art (sometimes referred to as its content)
- is the narrative, iconic, or nonobjective meaning conveyed by a figurative or an abstract composition.
- It is what is depicted in and by a work of art.
- It also covers the function of an object or architecture that otherwise has no narrative content.

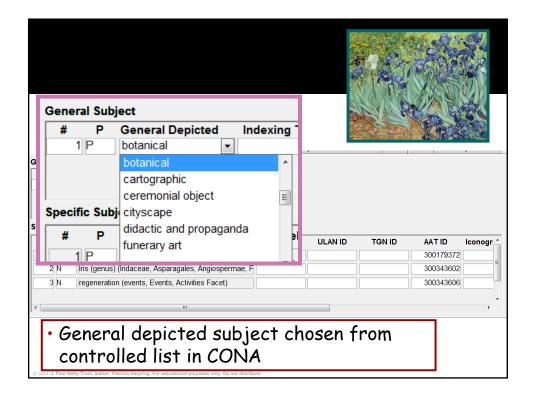
- In CONA (per CCO and CDWA) it is broadly defined
- "Core" information
- For *all* objects, even those that have no "subject" in the traditional sense

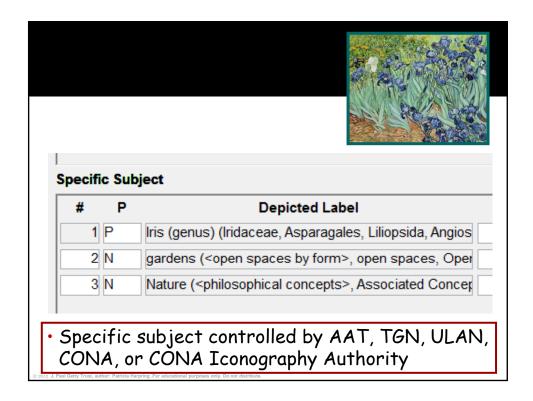
12 J. Paul Getty Trust, author: Patricia Harpring .For educational purposes only. Do not distribute

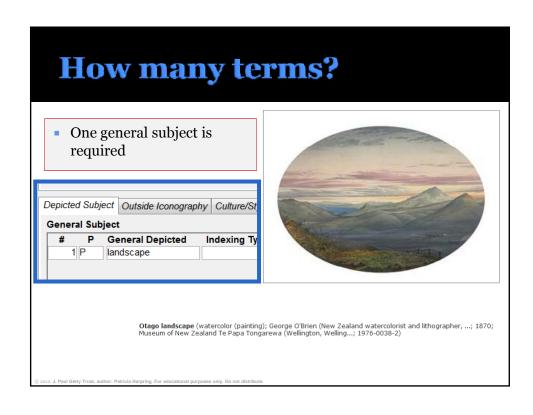
How many terms are required?

- Minimum requirement for CONA is one general subject term, although a default "undetermined" is available if absolutely necessary
- General subject in CONA is drawn from an extensible controlled list, suggested in CDWA and CCO
- Specific subject is controlled by authorities: ULAN, AAT, TGN, CONA, Iconography Authority
- Additional subject terms would allow better retrieval on subject, within the parameters of reasonable specificity and exhaustivity
- Extremely exhaustive subject indexing may not be worthwhile for projects not already doing so

2012 J. Paul Getty Trust, author: Patricia Harpring .For educational purposes only. Do not distribut







How many terms

 More are welcome for better access

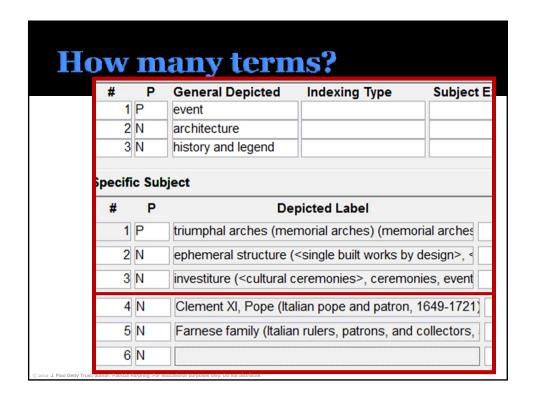
The print shows a triumphal arch erected in honor of Pope Clement XI's taking possession of San Giovanni in Laterano on April 10, 1701. Designed by Carlo Fontana and commissioned by Duke Francesco Farnese of Parma and Piacenza...

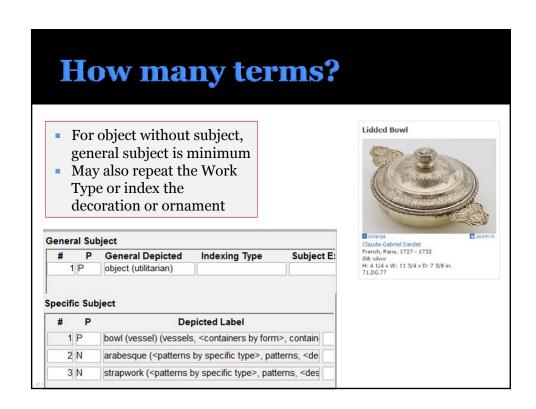


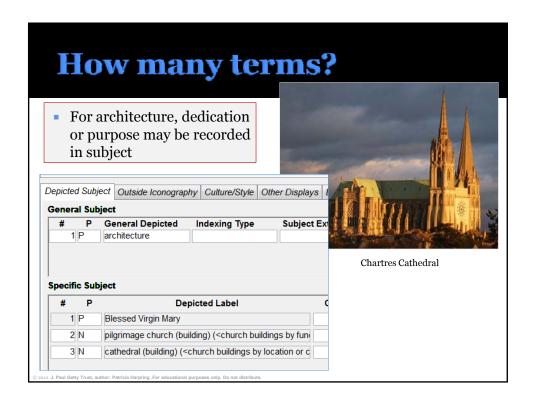
Arco trionfale fatto erigere dal ser.mo sig.r dvca di Parma alle glorie della santità di n.s. papa Clemente XI. in occasione del passaggio della Santità Sva al possesso di S. Gio. Laterano il dì X. aprile M.DCCI

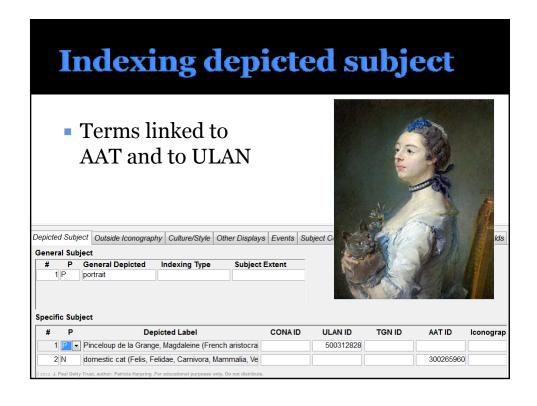
(etching (print); Specchi, Alessandro (etcher, 1668-1729); published: Nella stamp.a di Domenico de Rossi alla Pace ...; published 1701; Getty Research Institute, Special Collections (Los Angeles, Los Angeles county, California, United States); 640986; (CMalG)1366-954; accession number: P910002** (FF. 849)) (Movable Work) [700000455]

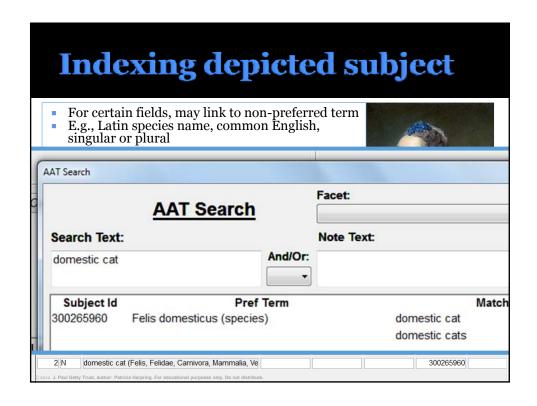
2012 J. Paul Getty Trust, author: Patricia Harpring .For educational purposes only. Do not distribute

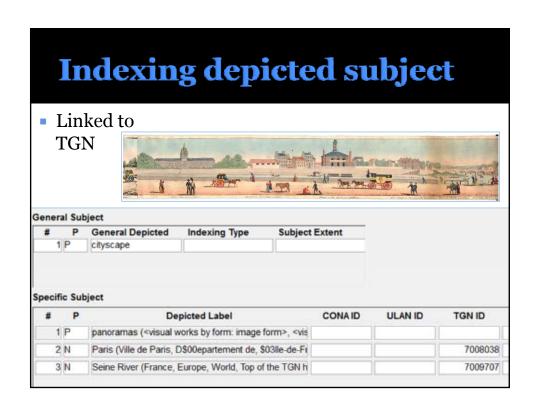


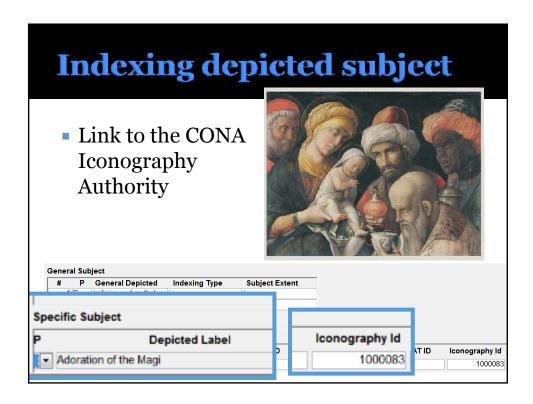








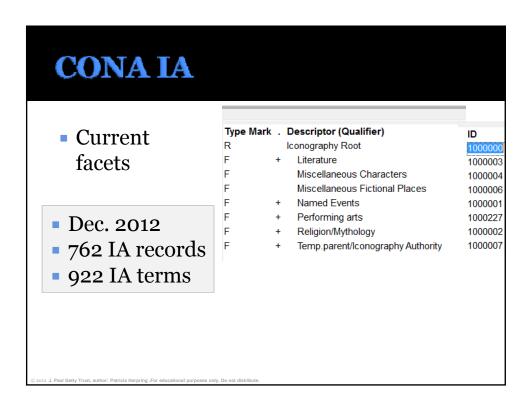


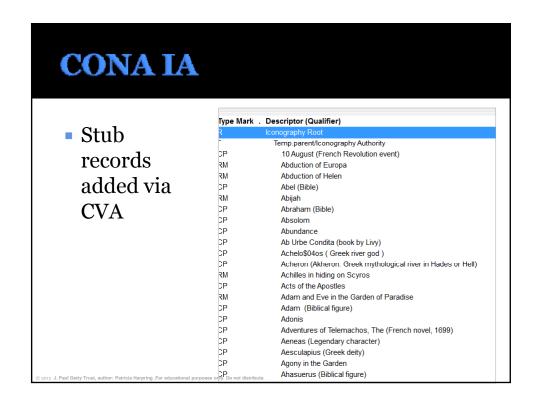


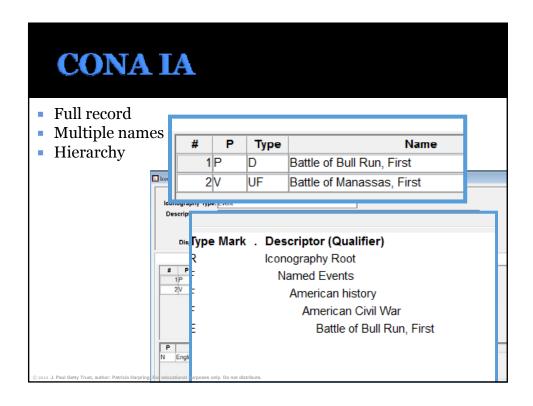
Indexing depicted subject

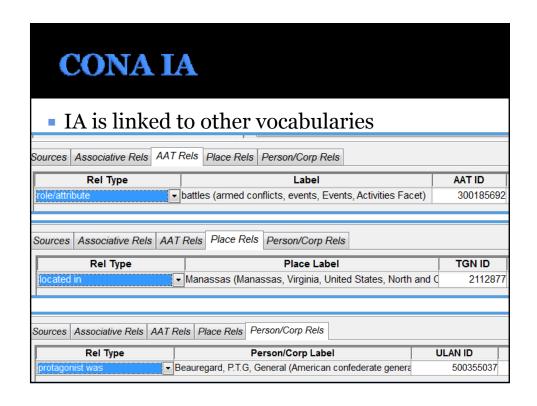
- The CONA Iconography Authority (IA), contains names/terms and other information for iconography and other subject terminology not contained in the other linked vocabularies
- The IA includes <u>proper names</u> for named events, religion/mythology, fictional characters and animals, themes from literature, and fictional places
- E.g., mythological and fictional places and characters (e.g., Zeus), the names of specific events, (e.g., the American Civil War), or iconographic themes (e.g., the Adoration of the Magi)

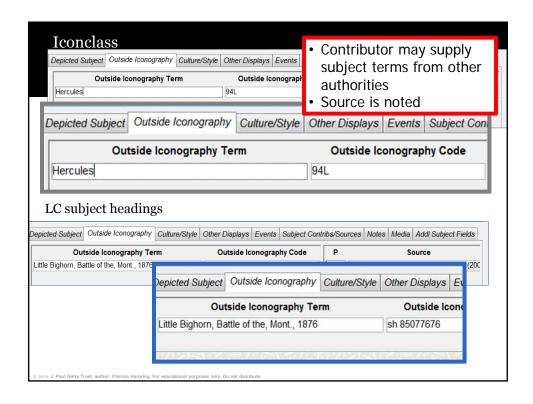
2012 J. Paul Getty Trust, author: Patricia Harpring .For educational purposes only. Do not distribute.

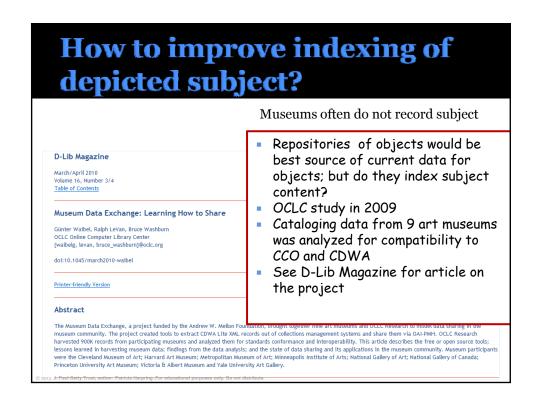












1. Work Type	0	86	86.5	89	82.5	85	78	100	86
2. Title	95.5	91.5	95.5	98	95.5	95.5	95.5	95.5	95
3. Display Creator	100	100	84	96	62	100	64	0	63
4. Name Creator	100	99	95.5	100	100	95	95	97	87
5. Nationality Creator	94	100	100	92.5	100	90.5	91.5	95.5	92
6. Vital Dates Creator	90	100	86.5	77.5	95	0	100	0	86
7 Role Creator	0	75	100	90	80	95	90	0	1
8. Display Measurements	100	100	100	91	100	100	100	100	77
9. Display Materials	100	100	100	95.5	100	100	100	100	10
Display Creation Date	100	91	100	95.5	100	95.5	100	100	8
11-12. Dates Concatenated (Earliest & Latest	99	91	95.5	95.5	99	100	95	100	
13. Location Name	100	99	77.5	100	100	80	99	80	97
14 WorkID	100	100	100	100	100	0	100	91	10
15. Subject Term	0	0	0	0	0	0	0	92.5	
17. Record Type	0	100	30	100	100	100	100	0	10

- Results of analysis show the percentages of records that have correctly included data for core CCO/CDWA fields
- Most museums generally are compliant with most of the core fields
- Some are missing data due to incorrect parsing (e.g., work type, role creator, which can be extrapolated and inserted globally in export)
- Exception: Only one of the nine museums had subject indexing

D12 J. Paul Getty Trust, author: Patricia Harpring .For educational purposes only. Do not distribute.

How to improve indexing of depicted subject?

- Subjects
 - Study results, subject often missing in museum records
 - In CONA contributions only occasionally included by museums
 - How to add it, from title, etc.
 - Extremely specific not useful examples of Iconclass and British museums effort

2012 J. Paul Getty Trust, author: Patricia Harpring .For educational purposes only. Do not distribute

How to improve indexing of depicted subject?

Some CONA contributors have provided some subjects

```
</PHYSICAL_CHARACTERISTICS>
- <DEPICTED_SUBJECT>
- <General_subject>
    <Person>Alphen, Simon van</Person>
    </General_subject>
    </DEPICTED_SUBJECT>
```

- How to improve subject indexing in contributions in the long run?
- In the immediate future, we are keen to accept contributions even without subject indexing

2012 J. Paul Getty Trust, author: Patricia Harpring .For educational purposes only. Do not distribute.



Title: Modern Rome—Campo Vaccino Joseph Mallord William Turner (English, 1775–1851)

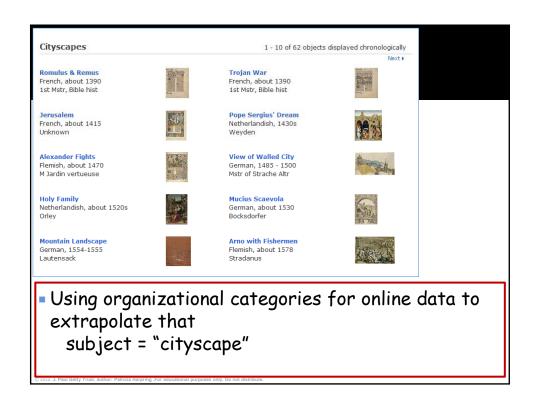
1839

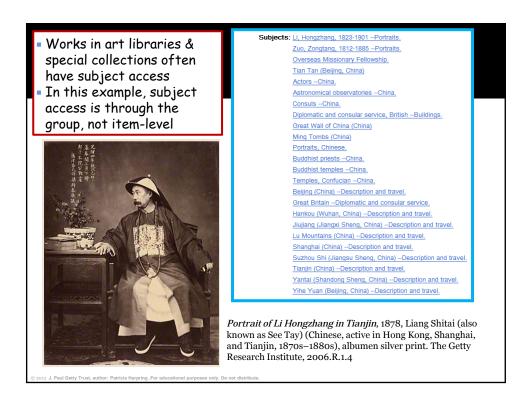
oil on canvas

36 1/8 x 48 1/4 in. (unframed), 48 1/4 x 60 3/8 x 4 3/8 (framed) J. Paul Getty Museum (Los Angeles, California); 2012.6

© 2012 J. Paul Getty Trust, author: Patricia Harpring .For educational purposes only. Do not distribute.

- Subject may often be found in title or work type
- but how to index it as subject?
- Automated parsing (matched against controlled subject lists) or entered by hand as subject?
- Combination of auto parsing and human editing would be best solution





What level of subject indexing?

- Three levels of analysis ("indexing type" in CONA)
 - Objective **description**

Of-ness

₹About-ness

- Identification of named subject
- Interpretation of the meaning or theme

Loosely based on method prescribed by Erwin Panofsky

- Editors: Never index beyond what is provided by the contributor or evident in the associated documentation
- CONA goal is minimal indexing and basicaccess indexing

2 J. Paul Getty Trust, author: Patricia Harpring .For educational purposes only. Do not distribute.

What level of subject indexing?

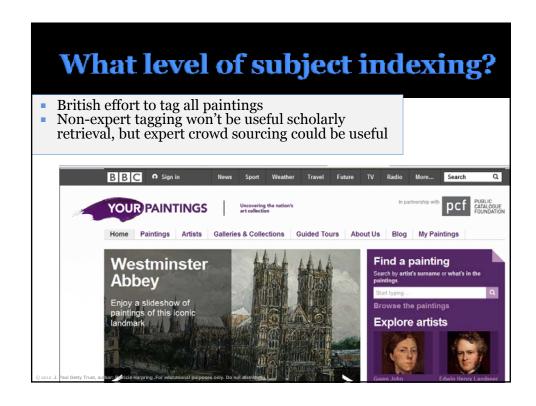
- Exhaustively detailed subject indexing is not required because that level of detail is not reflected in overall database thus will not aid overall retrieval
- Bottom line: if you are already indexing at detailed level, great. If not, you may want to strategize best way to use time for minimal indexing over entire collection
- Many Iconclass designations may be more detailed than required
- Example: Probably Apollo and Daphne would be enough

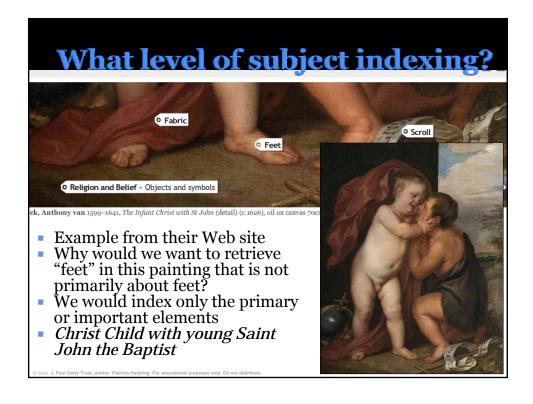
92B325 Apollo pursuing the nymph Daphne (as yet not metamorphosed)

97AA1 Daphne changed into a laurel-tree: while fleeing from Apollo,
Daphne is changed into a laurel tree

92K1461 The river-gods condole with Peneus for the loss of Daphne transformed into a tree

2012 J. Paul Getty Trust, author: Patricia Harpring .For educational purposes only. Do not distribute





Outstanding Issues

Editorial and technical issues to be resolved

2012 J. Paul Getty Trust, author: Patricia Harpring .For educational purposes only. Do not distribute

List of issues that were discussed at ITWG meeting 2013

- Terms associated with IDs changed
- How to incorporate data changes in "round trip"
- LOD vs existing formats
- No descriptor for a language
- Legacy terms in various languages
- True synonyms, exact equivalence
- Should microthesauri be contributed?
- Adding associative relationships
- Avoid unbound compound terms
- · Qualifiers are needed
- · Adding animal and plant species
- Other Specific questions
- · Controlled lists published as authorities
- How to flag variations or dialects of one language
- CONA Credit Line
- Linking non-IA entries to IA
- Include subject_id of linked vocab in contribution
- Other issues? Male/Female distinctions

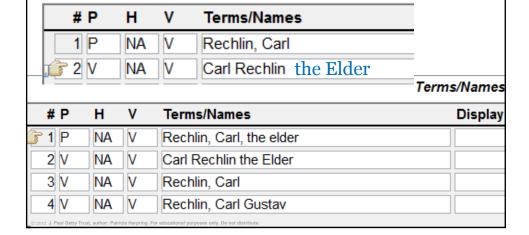
Terms associated with IDs changed

- For the past several years, editors have been very careful to avoid changing term_IDs, but was not done historically;
 - subject_id was persistent, but term_id could change (the term associated with it could change)
- But it is not enforced in the editorial system so it still happens sometimes
- Cannot be eliminated entirely, due to human error

2012 J. Paul Getty Trust, author: Patricia Harpring .For educational purposes only. Do not distribute

Terms associated with IDs changed

• Typical example where it happened in the past: display name in ULAN was edited instead of deleted and replaced



Terms associated with IDs changed

- But another significant problem still exists
- We do not track term "moved' to another record
- "Moved" conceptually, but actually deleted UF in one record is re-typed with a new term ID in a new record
- How to fix this in VCS?

- *great chairs* was UF in record for Harvard chairs, but not really an exact synonym for *Harvard chairs*
- A given *Harvard chair* may also happen to be a *great chair*, but not all *Harvard chairs* are *great chairs* and most *great chairs* are not *Harvard chairs*

Note: Armchairs designed to be the largest pieces of furniture in a sitting room, particularly common in 17thcentury Britain and America. They are made of wood, usually with conspicuous carvings and occasionally with upholstered seats. They may resemble a throne.

Terms:

great chairs (preferred,C,U,English-P,D,U,PN) (Dutch,D,U,U) great chair (C,U,English,AD,U,SN)

(Dutch,AD,U,U)

chairs, great (C,U,English,UF,U,N)

great-chairs (C,U,English,UF,U,U)

great armchairs (C,U,English,UF,U,N)



Note: Chairs having a triangular seat and three turned legs, popular in 17th-century America and Britain. The seat may be upholstered. The type is heavily inspired by Late Gothic and Early Renaissance three-cornered chairs of continental Europe.

Terms

Harvard chairs (preferred,C,U,English-P,D,U,PN) Harvard chair (C,U,English,AD,U,SN) chairs, Harvard (C,U,English,UF,U,U) sillas Harvard (C,U,Spanish-P,D,U,PN) silla Harvard (C,U,Spanish,D,U,SN)

Terms associated with IDs changed

- The Revision History notes that a term was deleted from the Harvard chairs record, but there is no link and no maintained ID when great chairs was re-typed into the new record
- Can we fix this? Requires new functionality in editorial system



AAT: Guide terms become concepts

• Some Guide Term records may become Concepts, as happens when users request it, e.g., *corner*

buttresses

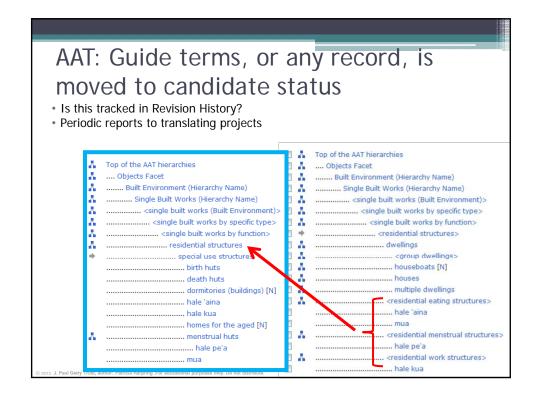
Subject ID	Descriptor (Qualifier)
300000763	<districts by="" condition=""></districts>
300000768	<administrative bodies=""></administrative>
300000796	<religious districts=""></religious>
300000900	<corner buttresses=""></corner>
300000955	<pre><piers by="" form=""></piers></pre>
300000966	<pre><piers by="" context="" location="" or=""></piers></pre>
300000955	<pre><piers by="" form=""></piers></pre>

2012 J. Paul Getty Trust, author: Patricia Harpring .For educational purposes only. Do not distribute.

Tracking of Guide Terms records that change *record type*

- GT becomes concept, GT becomes a candidate in a section of AAT called "to be deleted"
- How is reflected in revision history
- How can this be tracked in a report for translating projects?





Big Issue: How to track changes

- Updates in the Round Trip, and Version Control
- Additions work well, new subjects and new terms or other data in existing subjects
- But updating changes and deletions is currently done by hand or not at all
- Updating contributors' data coming to AAT:
 - How can we do this? Comparison of all fields in a new contribution with previous contribution may not work
 - Focus on a few fields? Which fields? SN, Term, Sources?

Big Issue: How to track changes

- Contributors updating AAT data coming to them:
 - Could you replace all data that is not from you as contributor?
 - For any data with you as contributor, you should be alerted of any changes, via reports
 - But most will want to see changes to English information too, or other languages
 - Again, could we target certain fields? Give you reports, or suggest reports you could write to compare your version to the new one?
 - Could we learn from what Gallery Systems is doing for AAT updates?

LOD vs existing formats

- Given user feedback and out analysis, it becomes clear that when we release the data as LOD, we cannot eliminate XML and Relational Tables
- While we cannot completely dismiss the possibility, if we ever were to discontinue them, we would give you much advance notice

No Descriptor for a language

- How does this happen? Legacy data
- Source (e.g., *IFLA Glossary*) listed only singular, which is the Alternate Descriptor
- Need expert contributor to add the plural
- AAT and all vocabs are works in progress

Legacy terms in various languages

- Expert approval needed
- RKD has already taken care of this for us, through reports and changes indicated on spread sheets
- Are these German, Italian, French, and Swedish terms correct?

solander boxes (preferred, C, U, English-P, D, U, PN) solander box (C,U,English,AD,U,SN) boxes, solander (C,U,English,UF,U,N) solander cases (C,U,English,UF,U,N) solanders (boxes) (C,U,English,UF,U,N) solander (solander box) (C,U,English,UF,U,U) cassettedozen (C,U,Dutch-P,D,U,U) cassettedoos (C,U,Dutch,AD,U,U) solanderdozen (C,U,Dutch,UF,U,PN) solanderdoos (C,U,Dutch,UF,U,SN) overslagdoos (C,U,Dutch,UF,U,SN) overslagdozen (C,U,Dutch,UF,U,PN) boîte à archives (C,U,French,AD,U,SN) Sammelschachtel (C,U,German,AD,U,SN) Kapsel (C,U,German,UF,U,SN) scatola a forma di libro (C,U,Italian,AD,U,SN) caja solander (C,U,Spanish,AD,U,SN) caja en forma de libro (C,U,Spanish,UF,U,SN) kapsel i form av en bok (C,U,Swedish,AD,U,SN) dokumentask (C,U,Swedish,AD,U,SN)

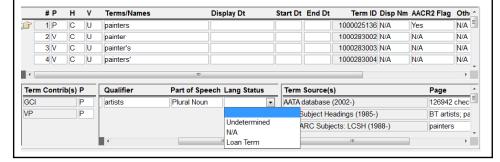
AAT: True synonyms vs not (): AAT Term Mapping Type Code 對應類型Mapping Type 適用情況Definition TELDAP 精準等同 【=】 The concepts of the two terms are exactly Exact, simple may need to equivalence ~EQ 近似等同 【><】 The concepts of the two terms are not exactly allow inexact the same. These concepts are equivalent in Inexact simple equivalence certain contexts while not in others, or these simple concepts might overlap in scope or are slightly different from each other in connotation equivalence 交集等同 【一對多】 The concepts of the two terms can be represented as two or more intersecting Intersecting compound • If so, we compound sets. equivalence 聯集等同 【一對多】 The concept of one term is equal to the sur need to flag Cumulative compound of the (two) concepts of another term. equivalence it in AAT 狹腐對應 【 < 】 The concepts of the two terms are in a narrower-to-broader hierarchical Hierarchical mapping data relationship. (narrower to broader) 廣狹對應 【>】 The concepts of the two terms are in a Hierarchical mapping broader-to-narrower hierarchical relationship. (broader to narrower) 關聯對應 The concepts of the two terms are neither equivalent nor hierarchically related, but are Associative mapping associative semantically to a certain degree.

AAT: True synonyms vs not

- We <u>strongly recommend</u> that translations are exact equivalences of the English descriptor
- Terms should be interchangeable
- All terms in the record should mean the same thing so that all terms are described by the SN and so that all Associative Relationships apply to the concept no matter what term is used to describe it

AAT: True synonyms vs not

- But given the difference in culture, perhaps it is necessary
- If you must do it, we want to capture Taiwanese flags for types for quasi-synonyms (*Exact* or *Inexact*; maybe also *Intersecting Compound*, etc.)
- Maybe add values to "Language Status," linked to language, which is linked to the term; should avoid changing data structure



Should Microthesauri be contributed?

- What should be contributed to the AAT?
- · Microthesaurus vs. new branch of the AAT
- If the terms are within scope of the AAT, and if they are not already in the AAT, <u>please</u> <u>contribute them</u>

Adding associative relationships

- When to add them?
- Only <u>unique</u> and critical relationships
- E.g., watercolor brushes are used only for watercolors; watercolorists paint only watercolors = unique relationship

When to make Associative Relationships

When to make Associative Relationships Make links to Related Concepts when it is useful to the end-user to have a cross-reference to the other concepts and when the concepts are directly related to each other but they do not have a hierarchical relationship. Think in terms of retrieval: Would such a link be useful in a search engine? If the concept records are clearly related but they are not linked via the hierarchy, create an associative relationship to prevent the conceptual link from being lost. In the examples below, the Related Concepts are in different hierarchies from the record for watercolors.

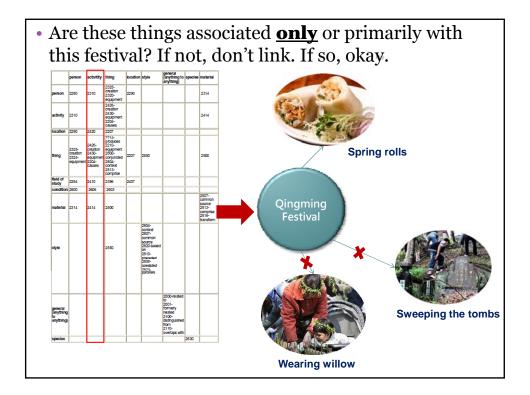
- Example
- [for "watercolors" paintings]
- Relationship Type: equipment needed/producing is Related Concept: watercolor brushes (<artists' brushes by function>, artists' brushes, ... Furnishings and Equipment)

Relationship Type: material(s) used/producing is Related Concept: watercolor (water-base paint, <paint by composition or origin>, ... Materials)

Relationship Type: material(s) used/producing is Related Concept: watercolor paper (drawing paper, <paper by function>, ... Materials)

Relationship Type: creator(s) is

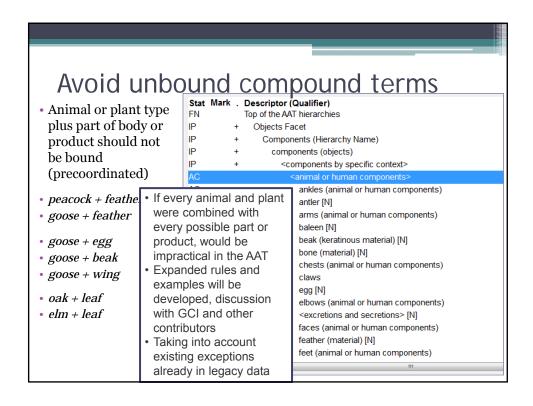
Related Concept: watercolorists (cpainters by technique, implement, or material employed>, implement, ... People)



Avoid unbound compound terms

- Contributions from GCI and elsewhere bring up renewed issues re. compound terms
- What are the criteria for precoordination in the thesaurus?
- When instead should the terms be postcoordinated, combined in search and retrieval?
- A compound term must express a single concept or unit of thought capable of being arranged in a genus-species relationship within the AAT hierarchy

- *flying buttresses* is bound (precoordinated)
- Style plus object type is not bound
- Baroque + cathedral
- Baroque + palace
- Baroque + painting
- Baroque + concerto
- Material plus object type is not bound
- bronze + sculpture
- bronze + bowl
- bronze + sword



Qualifiers are needed: clarification

- Qualifiers for one language term are required where there is a homograph for another language
- This is often omitted
- We can periodically give reports to contributors listing terms that need qualifiers in the mother AAT

Qualifiers needed: clarification

• *arca* is a homograph and needs a qualifier no matter what the language

```
    arcae (chests (case furniture), case furniture, ... Fur arca (safe chests)
    arks (case furniture) (case furniture, <storage and display furniture [300139717] arca</li>
```

Qualifiers needed: clarification

- arca is a homograph and needs qualifier
- Spanish, Dutch, and Italian terms need qualifier

```
arks (case furniture) (preferred,C,U,English-P,D,U,PN)
ark (case furniture) (C,U,English,AD,U,SN)
    (kastmeubilair) (Dutch,AD,U,U)
ark chests (C,U,English,UF,U,N)
arken (C,U,Dutch,D,U,U)
arcas (C,U,Spanish-P,D,U,PN)
arca (C,U,Spanish,AD,U,SN)

arcae (preferred,C,U,English-P,D,U,PN)
    (Dutch,D,U,U)
arca (safe chests) (C,U,English,AD,U,SN)
    (Dutch,AD,U,SN)
```

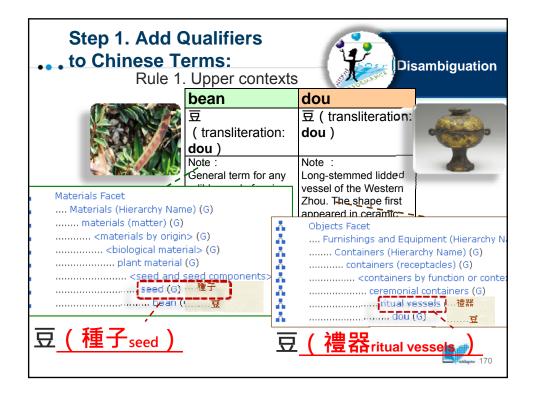
(Italian, AD, U, SN)

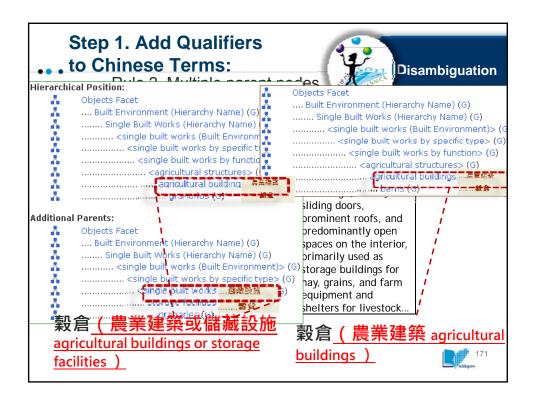
Qualifiers needed or not

• Taiwan project steps for creating qualifiers: Yes, this is compliant with AAT rules

Disambiguation – Rules for ... creating qualifiers

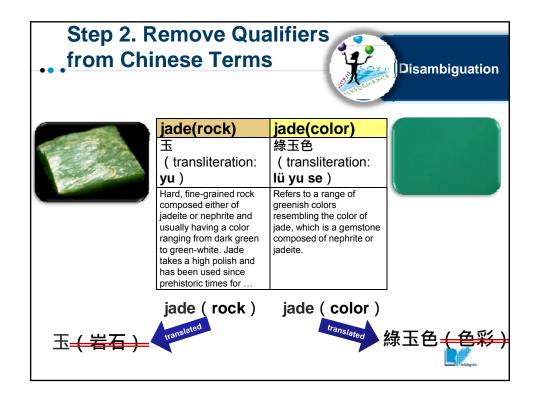
- Step 1. Add Qualifiers to Chinese Terms
 - Chinese qualifiers are added according to the priority rules below:
 - Upper contexts*
 - Multiple parent nodes*
 - · Special arrangement
- Step 2. Remove Qualifiers from Chinese Terms
 - Chinese qualifiers are removed according to the elimination of homograph







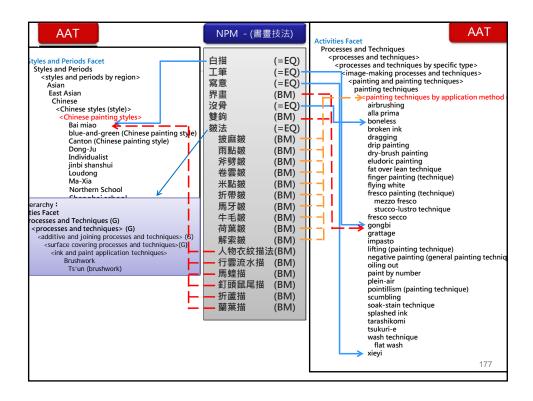




A few questions about specific additions or links

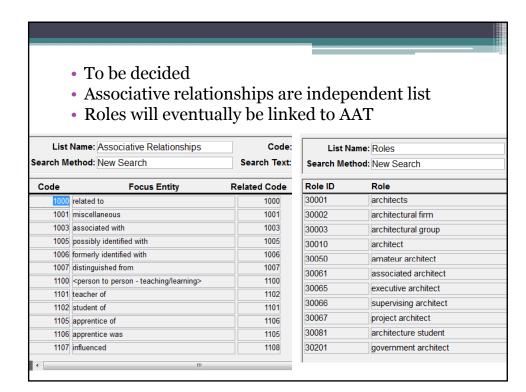
The following issues were discussed and resolved

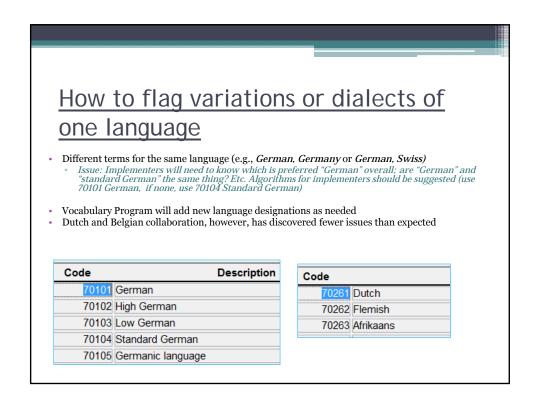


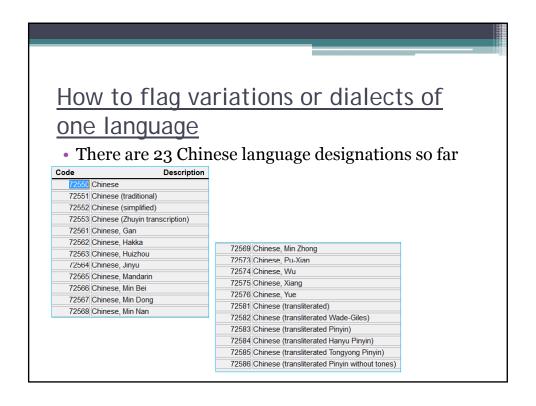


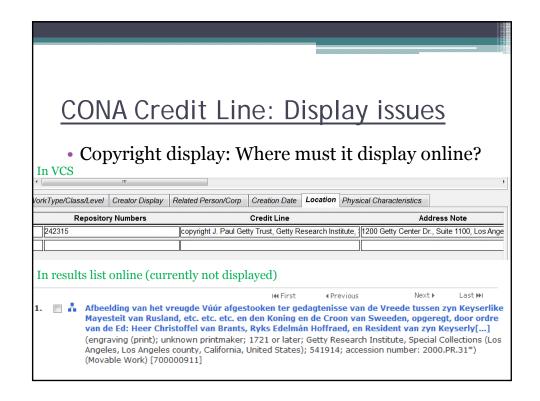
Various Vocab controlled lists as published authorities

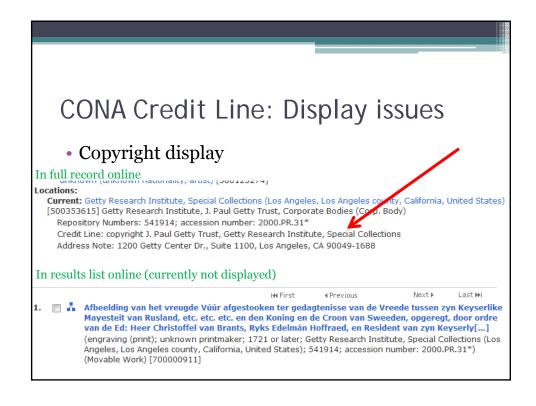
- Making controlled lists in vocabularies authorities in their own right
- Advantage is that others can use them in triples and elsewhere, Getty maintains brand on them
- When the list will eventually be linked to AAT, etc.; even so, this subset of the AAT has value as a separate thing
- When the controlled list (relationship types) is not directly link-able to another vocabulary

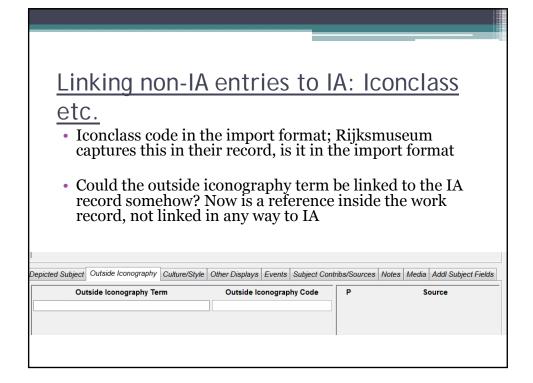












```
Include subject_id of linked vocab
  record if you know it
Object type
"panel
                   --<record>¶
painting" is
                   ··<Record Type>Movable Work</Record Type> • ¶
                   • < Catalog Level > Item < / Catalog Level > ¶
translated
                   °·<Object_Type>schilderij</Object_Type>·¶
from the
                   Classification
                     TITLES> 9
Dutch AAT
                   ··<Title>De aanbidding der koningen</Title> ¶
(subject_id?)
                   ° · < Preference > preferred < / Preference > • ¶
                   °·<Type>repository</Type>·¶

    Artist linked

                   •·<Language>Dutch</Language>
to ULAN
                   • < Title > The adoration of the Magi < / Title > ¶
                   °·<Preference>variant/alternate</preference>·¶
•Etc.
                    <Language>English</Language> ¶
                     ··</TITLES>¶
                   --<CREATOR>
                   • < Creator Display> Jan-Jansz-Mostaert (Haarlem-1474-- Haarlem-
                   1553)</creator_Display> ¶
•-<Corp.body_link>Mostaert,Jan Jansz</corp.body_link> ¶
                   °-<Role>schilder</Role>-¶
```

Others issues? • Male/Female designations: • AAT, as originally designed monolingual, inconsistently placed male and female roles in the same or separate records (duke / duchess) • Going forward with translations, this becomes a big problem, because the distinction between male and female will be different in different languages • Resolved with qualifiers? • No. Better solution is to eventually merge existing English male and female records into single records, to be considered synonyms rather than related concepts • Also better for retrieval if they are in the same records • Which is preferred, male or female? Is the other an AD? duchesses (noblewomen) (noblewomen, nobles (aristocrats), ... People) [300156290] duchesses' (noblewomen) duchesses' (noblewomen) duchesses' (noblewomen) duchesses' (noblewomen) duchesses' (noblewomen)

