Getty Vocabulary Program: Update
ITWG 2020

Patricia Harpring
Robin Johnson
Jon Ward
Antonio Beecroft

6 February 2019
Getty Vocabularies General Update:
Then, Now, Future
background, statistics, new activities

Patricia Harpring, Managing Editor
Getty Vocabulary Program
Getty Research Institute
What Are the Getty Vocabularies?

- AAT, TGN, ULAN, CONA, and IA contain structured terminology
- Thesauri, 3 relationships: equivalence (synonyms), hierarchical (broader-narrower), associative (btwn records, not hierarchical)
- Standards-compliant: comply with standards for controlled vocabularies (ISO, NISO); technical, releases also compliant
- Collaborative: Getty Vocabularies grow through contributions from expert user community -- institutions, consortia, and projects (300); contributing institutions share terminology gathered in the course of their own work, or through translation projects, for the benefit of the community
- Authoritative: Vocabularies cite authoritative sources and contributors
- Scope: Vocabularies strive to be ever more multilingual, multicultural, and inclusive
  Traditional audience (art: visual resources, catalogers, museums), plus scholars, and related disciplines (archaeology, conservation)
Summary of Recent Activities

- Processing contributions and growing Vocabularies
  - AAT: Translations: Processing Chinese, Dutch, German, French. Other terms for conservation, new media, multicultural terminology
  - TGN: Ancient sites; a few modern hierarchies, upon which we build historical data
  - ULAN: Loading contributions women artists, IMJ artists, modern artists, German and Austrian artists; African American artists
  - CONA: Architectural records in pre-processing
  - IA: Pre-columbian iconography, with Iconclass, Buddhist iconography

- OpenRefine Reconciliation Service using Getty Vocabularies

- Outreach: Presenting and publishing in various venues, AB, RJ, JW, PH; JW vice-chair of the board ICOM/CIDOC

- Strategic planning at GRI

- GDi technical roadmap, applications using Vocabularies
Statistics in Context
History of Getty Vocabularies

- Looking back 20 years, as we begin a new decade*
- Thanks to early visionaries committed to art information access for research and cataloging, began in 1979
- Enormous growth in size and scope
- Global reach with international partners: **THANK YOU!**
- Scope = multicultural and multilingual
- From publishing data in hardcopy books and limited licensed formats to open access and linking (LOD)
- Incorporation in search engines & collections management systems, ever more opportunities in expanding technical environment

* 00 vs 01
Growth of the Getty Vocabularies

- Comparison of numbers of concepts and terms over the last two decades (linked to each concept may be multiple terms / names)

<table>
<thead>
<tr>
<th></th>
<th>1997 *</th>
<th>2010</th>
<th>2020</th>
</tr>
</thead>
<tbody>
<tr>
<td>concepts/terms</td>
<td></td>
<td>concepts/terms</td>
<td>concepts/terms</td>
</tr>
<tr>
<td>AAT</td>
<td>30,000/70,000</td>
<td>48,100/131,000</td>
<td>70,000/436,300</td>
</tr>
<tr>
<td>TGN</td>
<td>904,500/922,600</td>
<td>912,100/1,106,000</td>
<td>2,532,000/4,086,000</td>
</tr>
<tr>
<td>ULAN</td>
<td>93,800/206,500</td>
<td>209,850/293,000</td>
<td>313,850/935,800</td>
</tr>
<tr>
<td>CONA</td>
<td>0/0</td>
<td>0/0</td>
<td>6,750/13,400</td>
</tr>
<tr>
<td>IA</td>
<td>0/0</td>
<td>0/0</td>
<td>1,900/5,200</td>
</tr>
</tbody>
</table>

* Year that the Getty Center opened
Recent/current preprocessing, loads  \textit{(no. of concepts)}

AAT: Chinese translation, festival terms, Academia Sinica (7,610)
AAT: French/English terms CHIN/Nomenclature (15,000)
AAT: Dutch terms, RKD & Belgian (1,410)
AAT: German terms, IfM-SMB (5,300)
AAT: Spanish, Chilean indigenous terms, CDBP-SNPC TRP (760)
AAT: Hebrew terms, IMJ (nnnnn)
AAT: Fashion terms, Europeana (480)
AAT: Italian, architecture, Bib. dell'Ac., Mendrisio (440)
AAT: Italian terms, ICCD (10,000, in processing)
AAT: various, Snite Museum, Notre Dame (31)
TGN: China & Switzerland, NGA/NIMA (650,170)
TGN: archaeological sites, NRW, Bonn (310)
**ULAN: Israel Museum, Jerusalem \textbf{(10,175)}**
ULAN: NMWA, Washington (2,250)
ULAN: Provenance Index, artists, owners (46,500)
ULAN: Knoedler, Szemann archives, GRISC (30,300)
ULAN: recent NACO contributions, GRI Library
ULAN: ancient potters, etc., INHA (660)
ULAN: artists, Belvedere, Vienna (22,000)
ULAN: various, A&AAL-UO 48
ULAN: various, SNITE: 27
**CONA: architecture, BWR (70,000)**
# Examples of pending loads and selected contributions under discussion

<table>
<thead>
<tr>
<th>Vocabulary</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAT</td>
<td>Chinese terms, Traditional Chinese Architecture Project</td>
</tr>
<tr>
<td>AAT</td>
<td>Japanese terms, NMJH</td>
</tr>
<tr>
<td>AAT</td>
<td>French terms, European BEINFRAT</td>
</tr>
<tr>
<td>AAT</td>
<td>Portuguese terms, Pinacoteca de São Paulo/U of São Paulo</td>
</tr>
<tr>
<td>AAT</td>
<td>Arabic terms, Louvre Abu Dhabi</td>
</tr>
<tr>
<td>AAT</td>
<td>Croatian terms, University of Zagreb</td>
</tr>
<tr>
<td>AAT</td>
<td>Nahuatl terms, Florentine Codex (4,000)</td>
</tr>
<tr>
<td>AAT</td>
<td>bronze casting, CASTing, JPGM Conservation et al (500)</td>
</tr>
<tr>
<td>TGN</td>
<td>various nations, NGA/NIMA ( &gt;200 files)</td>
</tr>
<tr>
<td>TGN</td>
<td>various places, Native American, Heard Museum</td>
</tr>
<tr>
<td>ULAN</td>
<td>Japanese artists, TNRICP</td>
</tr>
<tr>
<td>ULAN</td>
<td>Native American artists, Heard Museum, Phoenix (26,000)</td>
</tr>
<tr>
<td>CONA</td>
<td>visual surrogates ?</td>
</tr>
<tr>
<td>IA</td>
<td>Aztec iconography, Florentine Codex</td>
</tr>
<tr>
<td>IA</td>
<td>association with Iconclass ?</td>
</tr>
<tr>
<td>IA</td>
<td>Buddhist iconography ?</td>
</tr>
</tbody>
</table>
• Selected additions are posted quarterly for users
• For new data files, see monthly data releases


The Getty Vocabularies
AAT, TGN, ULAN, CONA, IA

Newsletter: The Getty Vocabularies grow

This document is intended for the general information of Getty Vocabulary end users, demonstrating the additions to the Getty Vocabularies over a given period of time, including new records and previous.

AAT: Added October-December 2019
7,938 new records, examples below. (Total does not include new data incorporated in “merged” records such as the Chinese or French translations of existing terms or otherwise processed as additions to existing records; does not include records loaded but still having “candidate” status.)

300024379 | animal cages | animal husbandry containers, <containers by function or context>, ...
Objects Facet
300241994 | hallmarking | stamping (marking), <surface marking processes and techniques>, ...
Activities Facet
300242002 | strain cracking | cracking, surface or structural changes, ... Activities Facet
300249525 | animals (general term) | <animals by form>, Animalia (kingdom), ... Agents Facet
300262597 | nyais | <people in service occupations>, <people by occupation>, ... Agents Facet
300262621 | simars | vestments (ecclesiastical costume), liturgical costume, ... Objects Facet
300262697 | tobacco rasps | <equipment for smoking and tobacco use>, personal equipment, ...
Objects Facet
300262909 | revetments (icon covers) | cases (containers), <containers by form>, ... Objects Facet
300263098 | display easels | easels, support furniture, ... Objects Facet
300263099 | supplédions | cushions, soft furnishings, ... Objects Facet
300263101 | étagères dressoirs | <storage and display furniture>, <furniture by form or function>, ...
Objects Facet
Growth of Contributor Community

https://www.getty.edu/research/tools/vocabularies/contributors.html

Contributors to the Getty Vocabularies

The Getty vocabularies grow in large part through continued efforts and contributions from the expert user community. Below is a list of contributors to the Getty vocabularies. The list is extensible and subject to frequent updates.

<table>
<thead>
<tr>
<th>Brief Name</th>
<th>Full Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>AAT-AD</td>
</tr>
<tr>
<td></td>
<td>AAT Advisory Committee (Williamstown, Massachusetts)</td>
</tr>
<tr>
<td>AS</td>
<td>AS-Academia Sinica</td>
</tr>
<tr>
<td></td>
<td>Academia Sinica (Nankan, Taiwan)</td>
</tr>
<tr>
<td>AAL</td>
<td>AALStanford</td>
</tr>
<tr>
<td></td>
<td>Art &amp; Architecture Library, Stanford (Stanford, CA)</td>
</tr>
<tr>
<td>UTA</td>
<td>UTA</td>
</tr>
<tr>
<td></td>
<td>Art and Art History Department, University of Texas at Austin</td>
</tr>
</tbody>
</table>

Numbers of contributing institutions

<table>
<thead>
<tr>
<th>1997</th>
<th>2020</th>
</tr>
</thead>
<tbody>
<tr>
<td>23 in USA, Canada, UK, Italy</td>
<td>300 globally</td>
</tr>
</tbody>
</table>
Growth of Users

- Usage steadily grows over time
- When data available as LOD, APIs, at other sites (RKD, Taiwan, etc.) online-search copy-and-paste usage on our site drops somewhat
- But still remain the most visited resources at GRI every month

Growth of Getty Vocabularies user community

<table>
<thead>
<tr>
<th></th>
<th>average per month</th>
<th>1997</th>
<th>2010</th>
<th>2020</th>
</tr>
</thead>
<tbody>
<tr>
<td>Users of online-search</td>
<td></td>
<td>1,500 mo</td>
<td>12,100 mo</td>
<td>11,600 mo</td>
</tr>
<tr>
<td>Visits / total page views</td>
<td></td>
<td>6,000 / 20,000</td>
<td>31,800 / 140,000</td>
<td>28,700 / 170,000</td>
</tr>
<tr>
<td>Users of LOD</td>
<td></td>
<td>0</td>
<td>0</td>
<td>7,000</td>
</tr>
<tr>
<td>Visits / hits on LOD</td>
<td></td>
<td>0/0</td>
<td>0/0</td>
<td>20,500 / 2,900,000</td>
</tr>
<tr>
<td>Locations where users are located</td>
<td></td>
<td>3 nations</td>
<td>85 nations</td>
<td>100 nations</td>
</tr>
</tbody>
</table>
### Usage on GRI site

<table>
<thead>
<tr>
<th>Feature</th>
<th>Views</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAT search</td>
<td>17380</td>
</tr>
<tr>
<td>TGN search</td>
<td>12117</td>
</tr>
<tr>
<td>ULAN search</td>
<td>10907</td>
</tr>
<tr>
<td>Provenance search</td>
<td>4425</td>
</tr>
<tr>
<td>Research home</td>
<td>4311</td>
</tr>
<tr>
<td><strong>Vocab home</strong></td>
<td><strong>4165</strong></td>
</tr>
<tr>
<td>Research tools</td>
<td>3746</td>
</tr>
<tr>
<td>Library home</td>
<td>3015</td>
</tr>
<tr>
<td>Provenance home</td>
<td>2885</td>
</tr>
<tr>
<td>BHA search</td>
<td>2247</td>
</tr>
<tr>
<td><strong>CONA/IA search</strong></td>
<td><strong>2210</strong></td>
</tr>
<tr>
<td>Photo Archive</td>
<td>1570</td>
</tr>
<tr>
<td><strong>CDWA home</strong></td>
<td><strong>1055</strong></td>
</tr>
<tr>
<td>GRI Portal</td>
<td>1018</td>
</tr>
<tr>
<td>Bauhaus</td>
<td>969</td>
</tr>
</tbody>
</table>

**Top 15 GRI Page Views, November 2019**
## Usage of LOD releases

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Month</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unique visitors</td>
<td>3,990</td>
<td>3,954</td>
<td>4,439</td>
<td>4,567</td>
<td>9,009</td>
<td>8,537</td>
<td>5,283</td>
<td>15,826</td>
<td>9,026</td>
<td>5,251</td>
<td>4,600</td>
<td>4,620</td>
<td>79,102</td>
</tr>
<tr>
<td>Number of visits</td>
<td>11,733</td>
<td>10,827</td>
<td>11,733</td>
<td>13,155</td>
<td>26,632</td>
<td>27,478</td>
<td>18,156</td>
<td>56,284</td>
<td>23,889</td>
<td>15,412</td>
<td>15,497</td>
<td>15,123</td>
<td>245,919</td>
</tr>
<tr>
<td>Pages</td>
<td>1,322,167</td>
<td>994,686</td>
<td>910,193</td>
<td>1,190,836</td>
<td>1,164,260</td>
<td>2,605,747</td>
<td>1,384,780</td>
<td>1,800,772</td>
<td>2,223,145</td>
<td>3,090,488</td>
<td>8,461,928</td>
<td>7,849,274</td>
<td>32,998,276</td>
</tr>
<tr>
<td>Hits</td>
<td>1,470,187</td>
<td>1,124,275</td>
<td>1,067,689</td>
<td>1,358,900</td>
<td>1,376,214</td>
<td>2,773,705</td>
<td>1,617,029</td>
<td>2,022,622</td>
<td>2,437,841</td>
<td>3,349,368</td>
<td>8,617,574</td>
<td>8,001,555</td>
<td>35,216,959</td>
</tr>
<tr>
<td>Bandwidth</td>
<td>103.15 GB</td>
<td>105.60 GB</td>
<td>85.48 GB</td>
<td>91.36 GB</td>
<td>107.38 GB</td>
<td>100.35 GB</td>
<td>77.43 GB</td>
<td>90.76 GB</td>
<td>187.77 GB</td>
<td>117.34 GB</td>
<td>168.43 GB</td>
<td>191.50 GB</td>
<td>1426.55 GB</td>
</tr>
</tbody>
</table>
Improving the Process
What Do We Do?

Contributions are preprocessed, loaded, and processed

- The Getty vocabularies grow through contributions from the expert user community
- Compiled and disseminated by Getty Vocabulary Program (GRI) and Getty Digital
- Released in online search, XML, Relational Tables, Web Services, JSON, RDF, N3/Turtle, N-Triples
What Do We Do?

System loader reconciliation gives us reports for near matches

• Contributions arrive in various formats and states of compliance
• Pre-Load data reconciliation
• During Load, merging, integration
• Post-Load, editing as necessary

<table>
<thead>
<tr>
<th>Prov place name</th>
<th>Problem</th>
<th>Place name in ULAN -- from TGN</th>
<th>ULAN Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Kartause) Champmol</td>
<td>not found</td>
<td>Chartreuse de Champmol (Côte-d’Or, Bourgogne, France)</td>
<td>4151060737</td>
</tr>
<tr>
<td>?</td>
<td>not found</td>
<td>[leave blank -- no link]</td>
<td></td>
</tr>
<tr>
<td>Aix</td>
<td>multiple matches</td>
<td>Aix-en-Provence (Bouches-du-Rhône, Provence-Alpes-Côte d’Azur)</td>
<td>4151000064</td>
</tr>
<tr>
<td>Albany</td>
<td>multiple matches</td>
<td>Albany (Albany county, New York state, United States) (inhabited place)</td>
<td>56000001030</td>
</tr>
<tr>
<td>Alessandria</td>
<td>multiple matches</td>
<td>Alexandria (Al-Iskandar$iyah governorate, Urban, Egypt) (inhabited place)</td>
<td>11230000108</td>
</tr>
<tr>
<td>Algier</td>
<td>multiple matches</td>
<td>Algiers (El Djazaïr department, Algeria) (inhabited place)</td>
<td>1001000031</td>
</tr>
<tr>
<td>Algiers</td>
<td>multiple matches</td>
<td>Algiers (El Djazaïr department, Algeria) (inhabited place)</td>
<td>1001000031</td>
</tr>
<tr>
<td>Alsace</td>
<td>multiple matches</td>
<td>Alsace-Lorraine (Europe) (general region)</td>
<td>40000000002</td>
</tr>
<tr>
<td>America</td>
<td>a 1765-1771</td>
<td>not found</td>
<td>United States (North and Central America) (nation)</td>
</tr>
<tr>
<td>Amersfoort</td>
<td>multiple matches</td>
<td>Amersfoort (Utrecht province, Netherlands) (inhabited place)</td>
<td>4390000025</td>
</tr>
<tr>
<td>Amiens</td>
<td>multiple matches</td>
<td>Amiens (Somme, Picardy, France) (inhabited place)</td>
<td>4151000131</td>
</tr>
<tr>
<td>Amsterdam</td>
<td>not found</td>
<td>Amsterdam (North Holland, Netherlands) (inhabited place)</td>
<td>4390000029</td>
</tr>
<tr>
<td>Amsterdam</td>
<td>multiple matches</td>
<td>Amsterdam (North Holland, Netherlands) (inhabited place)</td>
<td>4390000029</td>
</tr>
<tr>
<td>Amsterdam (Noord-Holland)</td>
<td>not found</td>
<td>Amsterdam (North Holland, Netherlands) (inhabited place)</td>
<td>4390000029</td>
</tr>
<tr>
<td>Ancona</td>
<td>multiple matches</td>
<td>Ancona (Ancona province, Marches, Italy) (inhabited place)</td>
<td>4243000167</td>
</tr>
<tr>
<td>Anet</td>
<td>multiple matches</td>
<td>Anet (Eure-et-Loir, Centre region, France) (inhabited place)</td>
<td>4151000158</td>
</tr>
</tbody>
</table>
What Do We Do?

Finding matches, merging duplicates representing the same person

<table>
<thead>
<tr>
<th>ULAN ID</th>
<th>Matching name</th>
<th>Display biography</th>
<th>Roles (AAT-controlled)</th>
<th>Nationality (AAT-controlled)</th>
<th>Estimated birth date</th>
<th>Estimated death date</th>
</tr>
</thead>
<tbody>
<tr>
<td>500017409</td>
<td>Guardi, Francesco</td>
<td>Venetian painter, 1712-1793</td>
<td>artist painter landscapist</td>
<td>Venetian Italian</td>
<td>1712</td>
<td>1793</td>
</tr>
<tr>
<td>500312969</td>
<td>Guardi, Francesco</td>
<td>Italian noble, born 1514</td>
<td>noble</td>
<td>Italian</td>
<td>1514</td>
<td>1614</td>
</tr>
<tr>
<td>500412666</td>
<td>Guardi, Francesco</td>
<td>Italian artist, ca. 1712-1793</td>
<td>artist</td>
<td>Italian</td>
<td>1707</td>
<td>1793</td>
</tr>
</tbody>
</table>

- Contributions arrive in various formats and states of compliance
- Pre-Load data reconciliation
- During Load, merging, integration
- Post-Load, editing as necessary

automated and manual integration
ULAN record is a merged, homogenous entity

Contributors

**Guardi, Francesco** (Italian painter, 1712-1793)

**Names:**
- Guardi, Francesco (preferred, V, index, LC)
- Francesco Guardi (V, display)
- Gardi, Francesco (V)
- Garde, Francesco (V)
- Gardis, Francesco (V)
- Gauda, Francesco (V)
- Guada, Francesco (V)
- Guarde, Francesco (V)
- Guardet, Francesco (V)
- Gvardi, Franchesko (V)
- Guardi,François (V)
- Guardi, Francisco (V)

**Sources**
- Avery Authority files (1963-)
- Harald Szeemann papers, Finding Aid, GRI Special Collections (2011-)
- J. Paul Getty Museum, collections online (2000-)
- Library of Congress Authorities database (n.d.) n 50074509
- M. Knoedler & Co. records, Finding Aid, GRI Special Collections (2012-)
- Provenance Index Databases, Authority file (1985-)
- RILA/BHA (1975-2000)
- Witt Library, Authority files

Note: Venetian painter, during the first half of his career focused on decorations and church paintings; later became vedutista (view painter) of Venetian urban life.

Attribution to contributor and source is maintained, but merged, unified, and "preferred" is flagged.

Links back to sources or contributors.
Improving the Process
How to streamline the processes?

- **Clearing the bottleneck** from contribution to publication
- Loading (GDi), pre-processing/processing (GVP) = points of congestion in workflow
- Critical to maintain high quality and authoritativeness in Getty vocabularies; careful vetting and sources = our greatest strengths
- Need to balance the speed of data release with avoiding publishing in-flux and wrong information
Improving the Process
How to keep quality high, but publish promptly

- Choose contributors with authoritative data
- If possible XML (many spreadsheets)
- Completeness of contributions = faster publication

- Goal: 80% of contributed records should require no editorial work, other than placing in hierarchy
- GVP has taken steps to relieve “processing” bottleneck (loading is GDi)
- Creation of holding hierarchies, published w/IDs
- OpenRefine being used for preprocessing
- Technical support for preprocessing, editors taking control where we can
- New or improved editorial system on the horizon
- Possibly new solutions to assist contributors
Tracking candidates & unloaded files

- Numbers of candidates loaded, awaiting processing
- In “temp.parents,” not published
- Numbers of records or files awaiting loading

**AAT Totals:**
- Concepts/terms: 69,856 / 436,334
- Candidates: 15,032 *
- In load queue: 1000

**TGN Totals:**
- Concepts/terms: 2,531,895 / 4,086,081
- Candidates: 26,425 *
- In load queue: 200 files

**ULAN Totals:**
- Concepts/terms: 313,825 / 935,788
- Candidates: 1,262 *
- In load queue: 22,000

* AAT: Italian terms; TGN: not places, but cultures; ULAN: attribution statements, not people
Improving the Process

Holding hierarchies

• Releasing concepts in holding hierarchies
• E.g., “Unidentified Named…”
• IDs are available to users
• But should be used with caution, as they may be merged with existing records later
• We will show you example of AAT load, for which the most of the merging was done prior to publication
Improving the Process

Editors participating in more pre-processing

- OpenRefine Reconciliation service using the Getty Vocabularies
- OR used to normalize, clean up data (parse names, make consistent)
- Reconciliation with AAT for role, nationality
- Editors are in training, will train contributors
Going Forward
Into the future

- We strive to make Getty Vocabularies ever more multilingual, multicultural, and inclusive

- Creating rich, authoritative vocabularies
- In order to meet requirements of GRI, Getty, and the field

- We plan to analyze weak areas or lacunae
- To seek contributors and sources of open data to fill the gaps

- Also, to address other issues that will enable Getty Vocabularies to remain relevant in the scholarship of various art histories
- Including linking to other resources
- And considering possibilities of digital scholarship and rapid advancement of technology
Examples of Links and Integration

- **VIAF**: ongoing, ULAN, concordance of IDs from resources globally. IMJ provided IDs.
- **Internal integration**
  - **Provenance Index**: AAT, ULAN, TGN; issue = local authority for artists verbatim names.
  - **PhotoTech**: AAT, ULAN, TGN; ? CONA.
  - **TMS**: Gallery Systems, updating Getty Vocabs on board; or link in a new Web product.
  - **Ex Libris**: Alma, how to incorporate Getty Vocabs, to enable reconciliation.

---

**Increase Links, Integration**

- To strengthen Getty Vocabularies’ five-star LOD status. See Zeng (2016):
  
  [https://www.getty.edu/research/tools/vocabularies/zeng_vocabs_and_five_star_lod.pdf](https://www.getty.edu/research/tools/vocabularies/zeng_vocabs_and_five_star_lod.pdf)

- We need more links out.
- Also, suggested: Which links **in** do we track?

---

- **VIAF**
  - Virtual International Authority File
  - Example: [http://vocab.getty.edu/page/ulan/500115588](http://vocab.getty.edu/page/ulan/500115588)
**Multicultural ULAN:**

Records total: 313,850
Nat/cultures total linked: 362,063

*a person may have multiple nationalities, cultures*

**Partial list of Nationality/Culture/Ethnicity**

American 57,624
Canadian 4,812
European 212,286
Mexican, Central-South 11,652
Native American, First Nations 4,394

*possible 26,000 Heard Museum*

African American 548
African 2,588
Middle Eastern 5,600
Indian subcontinent 1,370
East Asian 6,270
Other 49,945
Multilingual terms in AAT
Records total: 69856
Terms total: 436,334
Languages on terms total: 461,926
[a term may link to multiple languages]

Selected languages with term count
English 176,899 [including loan terms]
Chinese 102,620
Dutch 63,341
Spanish 56,182
German 21,315
French 21,139
Italian 4,439
Portuguese 247
Greek 63
Hebrew 29 [pending IMJ]
Russian 20
Arabic 16
Japanese 14
Persian 7
Other 13,873

Total Languages with Terms

ITWG 2020: Getty Vocabularies Update
Distribution of users, Vocabs online search
149 countries total; top 50

<table>
<thead>
<tr>
<th>Country</th>
<th>Users</th>
</tr>
</thead>
<tbody>
<tr>
<td>United States</td>
<td>79597</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>9125</td>
</tr>
<tr>
<td>Canada</td>
<td>6528</td>
</tr>
<tr>
<td>Germany</td>
<td>5809</td>
</tr>
<tr>
<td>France</td>
<td>5687</td>
</tr>
<tr>
<td>Italy</td>
<td>4847</td>
</tr>
<tr>
<td>Spain</td>
<td>4343</td>
</tr>
<tr>
<td>Netherlands</td>
<td>4025</td>
</tr>
<tr>
<td>Australia</td>
<td>3378</td>
</tr>
<tr>
<td>Belgium</td>
<td>2106</td>
</tr>
<tr>
<td>Poland</td>
<td>2017</td>
</tr>
<tr>
<td>Switzerland</td>
<td>1962</td>
</tr>
<tr>
<td>Russia</td>
<td>1602</td>
</tr>
<tr>
<td>Japan</td>
<td>1572</td>
</tr>
<tr>
<td>Mexico</td>
<td>1543</td>
</tr>
<tr>
<td>Brazil</td>
<td>1398</td>
</tr>
<tr>
<td>China</td>
<td>1200</td>
</tr>
<tr>
<td>Greece</td>
<td>1200</td>
</tr>
<tr>
<td>Argentina</td>
<td>1188</td>
</tr>
<tr>
<td>Taiwan</td>
<td>1087</td>
</tr>
<tr>
<td>Israel</td>
<td>896</td>
</tr>
<tr>
<td>Austria</td>
<td>894</td>
</tr>
<tr>
<td>India</td>
<td>876</td>
</tr>
<tr>
<td>Hong Kong</td>
<td>721</td>
</tr>
<tr>
<td>Philippines</td>
<td>667</td>
</tr>
<tr>
<td>Colombia</td>
<td>612</td>
</tr>
<tr>
<td>New Zealand</td>
<td>594</td>
</tr>
<tr>
<td>South Korea</td>
<td>565</td>
</tr>
<tr>
<td>Portugal</td>
<td>553</td>
</tr>
<tr>
<td>Sweden</td>
<td>542</td>
</tr>
<tr>
<td>Czech Republic</td>
<td>522</td>
</tr>
<tr>
<td>Ireland</td>
<td>480</td>
</tr>
<tr>
<td>Ukraine</td>
<td>468</td>
</tr>
<tr>
<td>Hungary</td>
<td>443</td>
</tr>
<tr>
<td>Slovakia</td>
<td>348</td>
</tr>
<tr>
<td>Turkey</td>
<td>348</td>
</tr>
<tr>
<td>Chile</td>
<td>336</td>
</tr>
<tr>
<td>Indonesia</td>
<td>330</td>
</tr>
<tr>
<td>Denmark</td>
<td>324</td>
</tr>
<tr>
<td>Croatia</td>
<td>306</td>
</tr>
<tr>
<td>Finland</td>
<td>300</td>
</tr>
<tr>
<td>Norway</td>
<td>294</td>
</tr>
<tr>
<td>Singapore</td>
<td>283</td>
</tr>
<tr>
<td>Puerto Rico</td>
<td>270</td>
</tr>
<tr>
<td>Egypt</td>
<td>246</td>
</tr>
<tr>
<td>Bulgaria</td>
<td>234</td>
</tr>
<tr>
<td>unidentified</td>
<td>222</td>
</tr>
<tr>
<td>South Africa</td>
<td>215</td>
</tr>
<tr>
<td>Thailand</td>
<td>193</td>
</tr>
<tr>
<td>Saudi Arabia</td>
<td>168</td>
</tr>
</tbody>
</table>
Update:
Getty Iconography Authority (IA)
and Cultural Objects Name Authority (CONA)
Accessing IA and CONA

- CONA and IA available in online search
- Also accessible as XML files, refreshed monthly
- Service for full list of records as XML is here:
  - http://vocabsservices.getty.edu/CONAService.asmx
- Service for individual records as XML is here:
  - http://vocabsservices.getty.edu/CONAService.asmx?op=CONAGetSubjectWeb
IA Sample Record

IA, Getty Iconography Authority™
http://vocab.getty.edu/page/ia/901000085

ID: 901000085

Names:
Avalokiteshvara (pref, en) [5000191]
Avalokiteśvara [5001067]
अलोकितेश्वर (sa) [5001068]
དཔལ་ལྡན་གཟིགས་(bo) [5002308]
Spyan ras gzigs (tib) [5002308]

Qualifier: Buddhist bodhisattva

AAT links
role/characteristic is ... bodhisattva
symbolic attribute is ... lotus
culture/religion is ... Mahayana (Buddhism)
culture/religion is ... Theravada (Buddhism)

Assessing Relationships
associated with .... Krishna (Hindu iconography)
counterpart is .... Guanyin (Buddhist iconography)

Hierarchical Relationships
Top of the IA hierarchies
Legend, Religion, Mythology
Buddhist iconography
Buddhist characters
Avalokiteshvara (Buddhist iconography)

Note
The bodhisattva of infinite compassion and mercy; embodies the compassion of all Buddhas. This bodhisattva is portrayed in different cultures as either female or male...

Contributors & sources
[VP] Encyclopedia Britannica Online (2002-); Bowker, Oxford Dictionary of World Religions (1997); Huntington, Art of Ancient India (1985); LC: LC control no.: sh 85010492

Includes proper names, relationships, and dates for iconographical narratives, religious or fictional characters, historical events, names of literary works and performing art

Links to other resources
LOC, Iconclass

Getty Vocabularies Update: IA and CONA
Indexing Subject Matter

- Iconography Authority used to index subjects not included in other Getty Vocabularies

**General Subject:**
- religion and mythology  *General Subject ID 31801*

**Specific Subjects:**
- Xibalba (Maya underworld)  *IA 1000045*
- Baby Jaguar (Maya character)  *IA 901002211*
- man (male human)  *AAT 300025918*
- skeleton (animal component)  *AAT 30019178*
- ax (weapon)  *AAT 300024664*
- altar  *AAT 300003725*
- death  *AAT 300151836*
- celebration  *AAT 300152441*

- **description, identification, interpretation**
- What the work is “of” and “about” (ofness, aboutness, isness) *
- Common iconographic themes, stories, and characters

CONA ID 70000012  Class: Precolombian art
Current Location: Metropolitan Museum of Art (New York, New York, USA)
IA and Iconclass

- Relationship between IA and Iconclass is currently under discussion.
- Overlap, particularly for Western iconography; also, Iconclass maps to AAT.
- IA contains names or short phrases in hierarchical arrangement; associative relationships.
- Long descriptive phrases have been considered out of scope for the IA, although IA links to the code (73E31).

![Diagram of IA Hierarchies]

### Iconclass

- **1A**: Deity, God (in general) ~ Christian religion
- **1B**: the Holy Trinity, 'Trinitas coelestis'; Father, Son and Holy Ghost ~ Christian religion
- **1C**: God the Father
- **1D**: Christ
- **1E**: the Holy Ghost
- **1F**: the Virgin Mary
- **1G**: angels
- **1H**: saints
- **1I**: prophets, sibyls, evangelists, Doctors of the Church; persons ~ the Bible (not in biblical context)
- **1K**: devil(s) and demons
- **1L**: Christian doctrine

- **1M**: 'Sapiencia' and the Seven Virtues, i.e. the Three Theological and the Four Cardinal Virtues.
CONA Sample Record

CONA, the Cultural Objects Name Authority®
http://vocab.getty.edu/page/cona/700001650

ID: 700001650

Catalog level: item
Class: prints
Work Type: engraving (print) reproductive prints
Date: Published 1679
Creator: Sébastien Le Clerc (engraver, 1637-1714)
Materials: 1 print: engraving, b&w
Dimensions: plate mark 40.8 x 50.2 cm, on sheet 45.3 x 53.4 cm
Inscriptions: Arc de triomphe de Louis XIV a la porte St. Antoine Arcus triumphalis Ludouici Magni ad Portam Antonianam / S. Le Clerc sculps. 1679.

Location:
current: Getty Research Institute, Special Collections (Los Angeles, California)
Repository Numbers: 2853-663; 853663; accession number: 2009.PR.68*

Note: The ephemeral arch was designed by Claude Perrault to commemorate Louis XIV’s entry …

Contributors: VP, GRISC
Sources: GRI Special Collections databases (1983-)
ULANTGN AAT

CONA, the Cultural Objects Name Authority®
This is a full record
Minimal record contains 5 fields
Link to GRI collection
Possible implementations of CONA

To record works depicted in other works (photo archive, prints, etc.)

To record works that are lost or designed but never built

CONA ID 700001288
Classification: architecture
Work Type: triumphal arch
Title: Arc de Triomphe du Trône
Location: Paris (France)
Creator: based on designs by Charles Perrault (French architect, 1613-1688), built by Claude Perrault (French architect, physician, scientist, 1613-1688)
Date: begun 1670, demolished 1716
Materials: stone and plaster
Dimensions: unavailable
General subject: architecture
Type: isness
Specific Subject: Louis XIV
Extent: honoree
ephemeral structure Extent: purpose

CONA ID 70001650
Title: Arc de triomphe de Louis XIV à la porte St. Antoine
Location: Getty Research Institute, Special Collections (Los Angeles) 2009.PR.68*

CONA ID 700001650
Title: Arc de Triomphe élevé à la Barrière du Trône, en l’honneur de Louis XIV
Location: Bibliothèque nationale de France (Paris, Ile-de-France, France) VE-2160 (6)-BOITE FOL

LINKS

depicts

depicted in

depicts

depicted in
Possible implementations of CONA

To link multiples

<table>
<thead>
<tr>
<th>CONA ID 700008519</th>
</tr>
</thead>
<tbody>
<tr>
<td>Catalog Level: multiples</td>
</tr>
<tr>
<td>Classification: prints and drawings</td>
</tr>
<tr>
<td>Work Type: drypoints</td>
</tr>
<tr>
<td>Title: Christ Presented to the People</td>
</tr>
<tr>
<td>Creator: Rembrandt van Rijn (Dutch, 1606–1669)</td>
</tr>
<tr>
<td>Description: Rembrandt treated this large drypoint almost as a painting, making marked changes to the composition as he reconceived …</td>
</tr>
</tbody>
</table>

| State: 1 |
| State: 2 |
| State: 3 |
| State: 4 |
| State: 5 |
| State: 6 |
| State: 7 |
| State: 8 |

- Note that each item, in various repositories, is cataloged as a unique work
- Then linked to the state, a conceptual relationship (not a physical relationship)
To record the work depicted in the surrogate

- What is advantage to linking to CONA for work depicted for Photo Archives? PhotoTech, PHAROS?
- Access for all surrogates that depict the same work
- If the work is not already in CONA, the cataloger could make a minimum record on-the-fly, to link to CONA ID
- If later the CONA record is merged, IDs are tracked

http://vocab.getty.edu/page/cona/700009600
### Possible implementations of CONA

**CONA ID:** 700002923  
**Catalog Level:** item  
**Classification:** architecture  
**Work Type:** building complex | temple

**Title/Name**  
Temple of Heaven (en)  
Tian Tan (zh-tr) | T'ien-t'an (zh-tr) | 天坛 (zh)  
| 天壇 (zh) | 천단 (ko)

**Creator**  
(display) first built by Ming Chengzu (1360-1424); extended by Ming Shizong (1507-1567); renovated by Qianlong (1711-1799)

**Creation Date**  
(display): 1530, during Ming dynasty

**General Subject**  
architecture  
Type: isness

**Specific**  
Earth [TGN 7030638]  
Extent: dedication

**Current Location**  
Beijing (Beijing Shi (municipality), China) (nation)  
Address Note: 39°57′05″N 116°24′36″E

**Dimensions**  
teire complex: 237 hectares

**Mat & Tech**  
(display) wood and stone [wood AAT 300011914 | stone AAT 300011176]

**Style:** Chinese [AAT 300018322] | Taoist [AAT 300143666]

**Contributors**  
JPGM VP GRISC BWR

**Sources**  

---

**Merging information from multiple contributors**

- CONA ID uniquely identifies the work in CONA
- Links to other IDs, such as LOC, VIAF, BWR
- Minimal record contains 5 fields

---

**To record architectural works**

**Built Works Registry** (Columbia, ARTstor)

**BWR data is in the queue for loading**

---

**Getty Vocabularies Update: IA and CONA**
Thank you.
Getty Vocabularies: ULAN
National Museum of Women in the Arts Contribution Project

Robin Johnson
Editor, Getty Vocabulary Program
Getty Research Institute
National Museum of Women in the Arts (NMWA) Database

NMWA in Washington DC features over 4,500 works of art by over 1,000 women artists in a variety of styles and media, spanning from the 16th century to present day.

2,250 records with various issues:

• Missing Nationality, Roles, Dates
  Example: Nil Yalter (artist, b. 1938)

• Two names for one artist in a record;
  Example: Kogga (Kolbrun Björgolfsdottir)

• Two separate artists names in one record
  Example: Forbush, Ann P. and Katherine M.D. Lobo
OpenRefine: Transforming Messy Data, One Column at a Time

Standardizing a Field in OpenRefine:

*Example: Role*

- Mapping Roles to terms in AAT Using Getty Vocabulary Reconciliation Service
- Separate multiple strings of roles in a single column into separate columns
- Uses default Role “artist” where role is blank
Loading NMWA Records into VCS

*Example:*

*Abakanowicz, Magdalena* (Polish textile artist, sculptor, 1930-2017)

- NMWA record merged with ULAN record (if one already exists)
- Fields, such as roles, went from being all in one Excel field, to being verified in OpenRefine, to VCS

<table>
<thead>
<tr>
<th>Role</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>artist</td>
<td></td>
</tr>
<tr>
<td>weaver</td>
<td></td>
</tr>
<tr>
<td>textile artist</td>
<td></td>
</tr>
<tr>
<td>sculptor</td>
<td></td>
</tr>
<tr>
<td>painter</td>
<td></td>
</tr>
</tbody>
</table>
Associative Relationships in ULAN: Meaningful links between Artists and Corporate Bodies

**Note:** Architect and designer, and chief of interior design for Artek. Aalto’s "Bögelblick" design serves as inspiration for a line of dinnerware produced by the Finnish decor company Iittala. Spouse and frequent collaborator of Alvar Aalto.

**Related People or Corporate Bodies:** collaborated with: Aalto, Alvar
related to: Artek
Thank you.

Kirsten Justesen
(Danish performance artist, born 1943)

*Lunch for a Landscape* (1975)

ULAN_ID 500112466

Getty :Vocabularies: NMWA contribution to ULAN
TGN In Focus: The Ancient World

Jonathan Ward
Editor, Getty Vocabulary Program
TGN In Focus: The Ancient World

→ **Over 10,000 archaeological sites**
  Includes places described as ruins, former settlements, rock shelters, and otherwise ancient sites.

→ **Over 400 ancient kingdoms and empires**
  Focus on African and Asian empires; Pre-Columbian, Native American / First Nations sites.
  
  ◆ **Focus on: places ignored by geodata systems**
    Various satellite and geodata systems bypass historic places entirely.

  ◆ **Focus on: value-added data for scholars / historians**
    Associative relationships; hierarchical relationships between historic places; linking a cultural object to a place in history.
TGN In Focus – Creating Links

1. Adel (former nation/state/emprise) (World, Africa) [5000047]
   Adel Kingdom

2. Ajarun Sultanate (former nation/state/emprise) (World, Africa, Somalia) [6698189]
   Ajarun

3. Aksumu (former nation/state/emprise) (World, Africa) [8711722]

4. Alwa (former nation/state/emprise) (World, Africa, Nubia) [5000219]
   Alodia

5. Anziku, Kingdom of (former nation/state/emprise) (World, Africa, Congo) [8698248]
   Kingdom of Anziku
   Teke
   Teke, Kingdom of
   Tio kingdom
   Tyo
   Tyo, Kingdom of

6. Asante (former nation/state/emprise) (World, Africa, Ghana) [5000434]
   Asante
   Asante empire
   Asante empire

7. Aksumite Empire (former nation/state/emprise) (World, Africa, Ethiopia) [7024409]
   Vernacular: Aksum
   Aksum, Le Royaume d'
   Aksum
   Aksumite Empire
   Aksum, Kingdom of
   Axum
   Kingdom of Axum
   Kingdom of Axum
   Kingdom of Axum
   Kingdom of Axum

Names:
- Kush (preferred, C,V,display)
- Cush (C,V)
- Kingdom of Kush (C,V)
- Kush, Kingdom of (C,V)

Hierarchical Position:
- World (facet)
- .... Africa (continent) (P)
  ........ Nubia (general region) (P)
  ............ Kush (former nation/state/emprise) (P)

Place Types:
- former nation/state/emprise (preferred, C) ............ ca. 1440 BCE-300 CE
- historical region (C)
- national region (H)
- kingdom (H)
- Nubian (C)

Related geographic places:
- capital is .... Napata ........ (deserted settlement)
- ............ (World, Africa, Sudan, Northern) [8711647]
- ............ ca. 790-591 BCE
- capital of .... Meroë ........ (deserted settlement)
- ............ (World, Africa, Sudan, Nile) [7000850]
- ............ from 591 BCE
Hierarchical Position:
- World (facet)
  - Africa (continent) (P)
  - Egypt (former nation/state/empire) (P)

Additional Parents:
- World (facet)
  - Byzantine Empire (former nation/state/empire) (P,H)
  - Egypt (former nation/state/empire)
- World (facet)
  - Roman Empire (former nation/state/empire) (P,H)
  - Egypt (former nation/state/empire) from 30 BCE

Top of the TGN hierarchy (hierarchy root)
- World (facet)
- Africa (continent)
- Nubia (general region)
- Kush (former nation/state/empire)
  - Athar an Na`a (deserted settlement) [N]
  - Athar el Musawwarat as Safra (deserted settlement) [N]
  - Kerma (deserted settlement) [N]
  - Moroc (deserted settlement) [N]
  - Napata (deserted settlement) [N]
  - Alwa (former nation/state/empire)
  - Ballanah (area) [N]
  - Buhen (deserted settlement) [N]
  - Makuria (former nation/state/empire) [N]
  - Mirgissa (deserted settlement) [N]
  - Qustul (area) [N]
  - Semna (deserted settlement) [N]
  - Wawat (former nation/state/empire)
TGN In Focus – Current and Upcoming
Native American / First Nation sites – Florentine Codex Project

Kituwah (deserted settlement)

Coordinates:
Lat: 35.26.13 N  degrees minutes  Lat: 35.4370  decimal degrees
Long: 83.24.25 W  degrees minutes Long: -83.4971  decimal degrees

Note: Kituwah is the site of the Cherokees' most important town and spiritual/cultural center, and is considered by some to be the originating point of all Cherokees. Moundworks, farms, and a council house were known to exist here prior to the mid-1600s. By the 1860s, the land was taken by settlers and the town destroyed. In 1996, the land was purchased by Cherokee elders to preserve the site.

Names:
Kituwah (preferred C V)

Hierarchical Position:
World (facet)
    .... North and Central America (continent) (P)
    ........ United States (nation) (P)
    ........ North Carolina (state) (P)
    ........... Swain (county) (P)
    ............. Kituwah (deserted settlement) (P)

Place Types:
deserted settlement (preferred, C)
historic site (C)
mound (C)
Cherokee (C)

Names:
Nicoya (preferred C, V, Spanish, latitude)
Necoc Yaol (H, V, Nahual, U) ........... meaning "between two waters [rivers]"

Hierarchical Position:
World (facet)
    .... North and Central America (continent) (P)
    ........ Costa Rica (nation) (P)
    ........ Guanacaste (province) (P)
    ......... Nicoya (inhabited place) (P)
Investigate inclusion of public data sets
Sites like scholar-driven Pleiades offer free and open, downloadable archaeological data that can integrate / merge with TGN; ARIADNE project partners; Wikidata?

Continued outreach with Getty partners

- GCI – AATA submissions; Linked Conservation Data Florentine Codex project
- Continue with lost kingdoms/settlements/sites
Thank you.
CHIN, RCIP, AASLH, Nomenclature Contributions to the Art & Architecture Thesaurus

Antonio Beecroft
Data Standards Editor, Getty Vocabularies
Réseau canadien d'information sur le patrimoine (RCIP) and Canadian Heritage Information Network (CHIN) submitted approximately 15,000 records to the AAT, of which ~5000 were merged into existing records. These represent French/English equivalents, expanding the number of French terms in the AAT, and adding more granular content, particularly in the Objects facet of the AAT.
**CHIN/Nomenclature Contributions Process**

1. Broader hierarchical contexts of contributed terms are mapped from Nomenclature to AAT to make potential merges of individual concepts possible (the CHIN/Nomenclature hierarchical structure does not match AAT).

2. Data in Nomenclature records is matched to and merged with existing AAT concepts.

3. New concepts that are not merged are moved to <temporary alphabetical lists> in appropriate places in the AAT hierarchy.

4. Placement of new concepts in AAT is evaluated, and checked again for duplicates (where CHIN/Nomenclature English Descriptors diverge from the AAT English Descriptor). Comparing the Nomenclature hierarchy with the AAT, in around 70% of cases it seems that the anticipated placement is close, but 30% or more diverge and an editorial decision must be made.

5. Records in temporary hierarchies gain additional content and context (Scope Notes, Associative Relationships, plural Descriptors) and are moved to more specific places in the AAT.
Hierarchies were mapped from Nomenclature to AAT to ensure accurate merges (Nomenclature, AAT hierarchies do not match one-to-one).

Matching subjects required the broader concepts for terms for the purpose of mapping.

CHIN/Nomenclature broader contexts were mapped to AAT broader contexts (in a spreadsheet).

<table>
<thead>
<tr>
<th>CHIN/Nomenclature</th>
<th>AAT</th>
<th>CHIN/Nomenclature Contributions to AAT</th>
</tr>
</thead>
<tbody>
<tr>
<td>300238982 personal equipment</td>
<td>NA</td>
<td>300238982 personal equipment</td>
</tr>
<tr>
<td>300238982 personal equipment</td>
<td>NA</td>
<td>300238982 personal equipment</td>
</tr>
<tr>
<td>300238982 personal equipment</td>
<td>NA</td>
<td>300238982 personal equipment</td>
</tr>
</tbody>
</table>

CHIN/RCIP/Nomenclature Contributions to AAT
2. CHIN/Nomenclature records are matched to and merged with existing AAT concept records under matching broader concepts.
For terms that are new to AAT, many remain in a temporary alphabetical lists, until they are reviewed and incorporated into the logical structure of the AAT.
Duplicate reports match new records to homographs in the AAT. For CHIN/Nomenclature records, comparing a branch of hierarchy from a given parent key to another branch.

Not all homographs are duplicates. Editors determine what is, and is not, a potential merge.
Records in temporary hierarchies are reviewed, and additional content and context is added

Ongoing review and edits:

- Addition of English Descriptors in plural form.
- Addition of Scope Notes and sources.
- Records are moved to appropriate places in AAT and attendant adjustments where parent/child relationships must be reconciled is made.
- Assign Associative Relationships where appropriate.
Thank you.