Overview

• Background

• The Asch Collection

• Modeling Cultural and Scientific Heritage

• The ASCH Use Cases and Requirements

• The ASCH Model
Collections at the University of Göttingen

• About 30 scientific collections of different disciplines
Digitization of collection items

• In most cases there is little known about the items
  – Few use machine-readable data for the descriptions of their items
  – The data is only accessible inside the institutions hosting the collections
  – The condition of many of the items doesn’t allow access by users

=> Most of the items seem to be lost for users.

• Most of those who catalogue their data in a database
  – Use proprietary, „homemade“ software (excel, fileMaker, MySQL, etc.)
  – Non-standardized, „homemade“ metadata descriptions
  – No authorities or other controlled values

=> Most of the data is not reusable
How to start?

• We need standards for
  – well-structured reusable descriptions of the items
  – Digitization of the items
  – Preservation of the data

• To make collection items
  – Findable
  – Obtainable
  – Accessible
  – Usable

• Don’t begin at the whiteboard
  – Ask the holding institutions, what they really need
  – Ask the scientists what they want to know
  – Ask them all what they use
GEORG THOMAS VON ASCH AND THE ASCH COLLECTIONS
Georg Thomas von Asch

Portraiture of Georg Thomas von Asch (1729-1807);
painting by Kyrill Golovachevski;
oil on canvas, 86 x 70 cm;
donation of 1780
Göttingen Royal Academic Museum, 1773
“Woman of the Reindeer Tungus in summer garment”,
watercolored copper engraving;

in:
Gavrila Sarychev: Gavrilas Sarytschews’s Russisch-Kaiserlichen Generalmajors von der Flotte achtjährige Reise im nordöstlichen Sibirien, auf dem Eismeere und dem nordöstlichen Ozean, (1805-1815);
SUB Göttingen, Russica S 265
Institute of Social and Cultural Anthropology:
The Ethnographic Collection

Shaman's Costume (front), Tungus (Evenks), Siberia;
hight: 190 cm;
As 957
Department of Anatomy and Embryology, Center for Anatomy, University Medical Center Göttingen: The Skull Collection

Skull of a Yakut;
Kratzsch, Bartholomeus (????-????);
AIG.559
Geoscience Centre: Historical Collections

Graphite; Geoscience Museum; with original labeling – von Asch gives some information on the graphit’s use
Department of Archaeology: The Coin Collection

Medal with the profile of von Asch and the inscription „Liberator a peste“;
diameter: 53 mm;
weight: 61.15 g
Evidences
### Department of Art History and Art Collection

<table>
<thead>
<tr>
<th>DatenbankID</th>
<th>127028</th>
</tr>
</thead>
<tbody>
<tr>
<td>Museum / Sammlungsbereich</td>
<td>&gt; Göttingen &gt; Georg-August-Universität Göttingen &gt; Kunstsammlung &gt; Sammlung Asch</td>
</tr>
<tr>
<td>Inventarnummer</td>
<td>Asch D 9</td>
</tr>
<tr>
<td>Alte Inventarnummern / andere Nummern</td>
<td></td>
</tr>
<tr>
<td>Objektbezeichnung / Gegenstand</td>
<td>&gt; Graphik, Photographie* &gt; Druckgraphik &gt; Tiefdruck &gt; Kunferstich</td>
</tr>
<tr>
<td>Titel</td>
<td>Porträt Michael Denis</td>
</tr>
<tr>
<td>Kategorie</td>
<td>Grafik</td>
</tr>
<tr>
<td>Objekttyp</td>
<td>Bild</td>
</tr>
<tr>
<td>Status</td>
<td>Einzelobjekt</td>
</tr>
<tr>
<td>in Bearbeitung</td>
<td>Nein</td>
</tr>
<tr>
<td>Veröffentlichen auf Unisammlungen Göttingen</td>
<td>Nein</td>
</tr>
</tbody>
</table>
Museum of Zoology
Letter from Asch to Heyne 28th February 1780

P.S. vermutlich wird Hr. Prof. Murray ein den 18/29 Januar abgeferdigtes Päckchen mit Sämereyen aus der Krimm, so ein von hier nach dasigen Gegenden abgereißen Kaufmann Benoit mitgenommen, schon erhalten haben.

Prof. Murray was head of the Botanic Garden since 1769.
About the Projekt

• Purpose and Scope: developing a metadata model for contextualising heterogenous objects from collections

• Funded by Deutsche Forschungsgemeinschaft

• Duration: 01.09. 2014 – 31.08 2017

• Project Lead:
  – Göttingen State and University Library
  – Institute of Social and Cultural Anthropology at the University of Göttingen

• Contact
  – Susanne Al-Eryani (mailto: al-eryani(a)sub.uni-goettingen.de)
  – Gudrun Bucher (mailto: gudrun.bucher(a)sub.uni-goettingen.de)
  – Jürgen Dönitz (mailto: doenitz(a)sub.uni-goettingen.de)
MODELING CULTURAL AND SCIENTIFIC HERITAGE
Cultural and Scientific Heritage in the WWW

• The vision
  – A one-stop shop for access to millions of digital cultural and scientific objects
  – Interlinking the results of different projects, institutions and disciplines

• The need
  – well-structured consistent description of objects using metadata standards

• The reality
  – Metadata standards used by cultural and scientific institutions sometimes differ widely in their model and structures
  – Metadata descriptions are aligned to the requirements of specific projects, in-house rules, legacy data, community agreements, …

And most time it makes sense!
Some Metadata Descriptions are …

• Description of text resources
  – MARC (Machine-Readable Cataloging) or MODS (Metadata Object Description Schema)

• Description of the events in the lifecycle of an object
  – LIDO (Lightweight Information Describing Objects) or CIDOC-CRM (Conceptual Reference Model of the International Committee for Documentation of the International Council of Museums)

• Description of resource compilations
  – EAD (Encoded Archival Description) or METS (Metadata Encoding and Transmission Standard)

• Description of resources from natural science
  – Darwin Core or ABCD (Access to Biological Collection Data)
and some are…

- **contextualisation**
  - EDM (Europeana Data Model), W3CPROV, skos (Simple Knowledge Organization System)

- **Interlink**
  - Who
  - When
  - What
  - Where

- **Interlink with**
  - evidences

[http://commons.wikimedia.org/wiki/File:Latin_America_-_First_level_political_divisions.svg](http://commons.wikimedia.org/wiki/File:Latin_America_-_First_level_political_divisions.svg)
“Mega-Model“ or „Molok“

- One standard that fits them all
- Watched by the metadata police
- What happens:
  - „… to re-work Diane Ingmann’s maxim ‘there are no metadata police’, implementors will bend the metadata schemas for their own purposes.“
    (source: http://www.agi-imc.de/internet.nsf/88184949ab26922885256871006f3fa5/f106435e0fd9ffcc1c125699f002ddf31/$FILE/dubin_core.pdf)
  - or develop a new one

picture: http://xkcd.com/927/
“Meta-Model“ or Framework

• Uses the least common denominator to bring the data together
• Allows different types of metadata descriptions

• Use
  – Dublin Core Dumb-Down Principle: „… a principled way of viewing a complex metadata description through the lens of a simpler representation …“

  – RDF (Resource Description Framework): a most generic framework modelling the relations between resources
    • Mandatory: every described resource must be identified by a unique http-Identifier
    • Freedom: use whatever metadata standard you like as long as the terms are identified by http-URIs
Scope of the Asch-Project

- Developing a metadata framework to bring together metadata descriptions of different types

- Focusing on provenance information
  - W3CPROV as the backbone
  - align it with other standards (EDM, CIDOC-CRM)
  - align it with evidences

- re-use of domain-specific standards
  - Using domainspecific standards like LIDO, MODS, Darwin Core, etc.
  - Using the DCMI dumb-down principle to align them with the meta-level
  - Check how good they fit into the model
PREVIOUSLY ...
Proceeding

- Interviews and workshop
  - Curator of the collections in Göttingen
  - Scientists working on collection items
  - Metadata experts from different collections

- Purpose and scope
  - collect use cases
  - find out, what the researcher wants to know
  - find out in what context provenance information is used and needed

- Asch Framework shall be:
  - less initiated by applications
  - more initiated by research questions and use cases
### 31 Case Studies

<table>
<thead>
<tr>
<th>Case Study</th>
<th>Scenario</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CS 17</td>
<td>Percy is searching for a manuscript with an unknown title. All he knows is that this manuscript was once part of a volume that also included the Octavo manuscript. Therefore he searches for the Octavo manuscript and gets the information that the volume was dispersed at the beginning of the 18th century and the two parts were sold to different collectors so that the two manuscripts are no longer parts of the same volume. But because the description of the former volume still exists, and both parts are still interlinked with it, Percy gets information about the whereabouts of the manuscript he needs.</td>
<td></td>
</tr>
<tr>
<td>CS 18</td>
<td>Anna is working on a thesis on the focus of private art collections at the end of the 19th century. She wants to analyze to what extent the interests of collectors changed during this time and whether there were differences between German and French collectors. She is especially looking for artworks which were bought on the French art market for more than 50,000 Francs during this time. Therefore she is searching for information about collectors of this time in Germany and France and the pieces they collected. She needs to know where the collectors came from, what were their interests and what items did they collect when. She is also searching for documents (e.g. letters, contracts, auction catalogues) about the move of items from one collection to another during this time.</td>
<td></td>
</tr>
<tr>
<td>CS 19</td>
<td>Antra got the task to search for prehistoric art from Lithuania abroad and verify to what extent the items she finds are looted art. Therefore she needs information not only about the locations of the items and the current collections they are part of but also about the way the items came to be collected. Antra gets this information by searching for all items of the same class.</td>
<td></td>
</tr>
</tbody>
</table>

### Scenario

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenario 35</td>
<td>A user wants to reconstruct the former relations of an item (e.g. an is-part-of relation).</td>
</tr>
<tr>
<td>Scenario 3</td>
<td>A user needs to know when an item was related to another item (e.g. was part of another item).</td>
</tr>
<tr>
<td>Scenario 6</td>
<td>A user needs biographical information about life and work of an agent.</td>
</tr>
<tr>
<td>Scenario 29</td>
<td>A user needs to know what items were related to an agent during a certain time span.</td>
</tr>
<tr>
<td>Scenario 31</td>
<td>A user needs information about the time an event occurred.</td>
</tr>
<tr>
<td>Scenario 25</td>
<td>A user wants to know where an event happened.</td>
</tr>
<tr>
<td>Scenario 48</td>
<td>A user needs access to the evidences of statements.</td>
</tr>
<tr>
<td>Scenario 24</td>
<td>A user is searching for items of the same price related to acquisition events.</td>
</tr>
<tr>
<td>Scenario 13</td>
<td>Searching for an item a user uses a controlled vocabulary to get all items of the same class.</td>
</tr>
<tr>
<td>Scenario 1</td>
<td>A user wants to know more about the collection an item belongs or belonged to.</td>
</tr>
</tbody>
</table>
8 Use Cases

- UC 1 Information about resources
- UC 2 Identification of resources
- UC 3 Information about the history/lifecycle of resources
- UC 4 Change of use and reception of resources
- UC 5 Proof of information by evidence
- UC 6 Provenance of statements
- UC 7 Access to resources
- UC 8 Reuse of data
## 8 Use Cases
related to the case studies via scenarios

<table>
<thead>
<tr>
<th>Use Case</th>
<th>Scenario</th>
<th>Actor</th>
<th>Action</th>
<th>Requirement ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>UC 3 Information about the history/lifecycle of resources</td>
<td>Scenario 34</td>
<td>editor</td>
<td>Interlink items with more than one collection</td>
<td>[Requirement 51] Item descriptions must be assigned to 1-n collections.</td>
</tr>
<tr>
<td>UC 3 Information about the history/lifecycle of resources</td>
<td>Scenario 35</td>
<td>user</td>
<td>Browse from an item to the items it is/was related to.</td>
<td>[Requirement 9] Resources should be interlinked with 0-n other resources.</td>
</tr>
<tr>
<td>UC 4 Change of use and reception of resources</td>
<td>Scenario 56</td>
<td>editor</td>
<td>Interlink a statement with resources describing the context of this statement.</td>
<td>[Requirement 54] Statements must be linkable with resources (e.g. evidence) proving their reliability.</td>
</tr>
<tr>
<td>UC 4 Change of use and reception of resources</td>
<td>Scenario 57</td>
<td>editor</td>
<td>Interlink a statements with date information.</td>
<td>[Requirement 74] Statements must be linkable with resource descriptions.</td>
</tr>
<tr>
<td>UC 4 Change of use and reception of resources</td>
<td>Scenario 58</td>
<td>editor</td>
<td>Interlink a statement with the description of a place.</td>
<td>[Requirement 74] Statements must be linkable with resource descriptions.</td>
</tr>
<tr>
<td>UC 5 Proof of information by evidence</td>
<td>Scenario 36</td>
<td>user</td>
<td>Search for all resources related to an item.</td>
<td>[Requirement 13] Item descriptions should be interlinked with evidence.</td>
</tr>
<tr>
<td>UC 5 Proof of information by evidence</td>
<td>Scenario 37</td>
<td>user</td>
<td>Search for all resources related to an event.</td>
<td>[Requirement 28] Event descriptions should be interlinked with documents about the event. [Requirement 19] Evidence accessibility must be described.</td>
</tr>
</tbody>
</table>
## About 70 Requirements

<table>
<thead>
<tr>
<th>Requirement 25</th>
<th>The function/role of an agent during an event must be specified (e.g. creator, dealer, collector) using controlled values.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Requirement 26</td>
<td>An event in the lifecycle of an item must be related to 0-n places.</td>
</tr>
<tr>
<td>Requirement 27</td>
<td>An event in the lifecycle of an item must be related to 0-n date information.</td>
</tr>
<tr>
<td>Requirement 28</td>
<td>Event descriptions should be interlinked with documents about the event.</td>
</tr>
<tr>
<td>Requirement 29</td>
<td>Date and time indication must be machine readable.</td>
</tr>
<tr>
<td>Requirement 30</td>
<td>Date information must be compliant to datatype standards.</td>
</tr>
<tr>
<td>Requirement 31</td>
<td>A human readable form of data information should only be used in addition to a machine readable form and should use values of a controlled vocabulary.</td>
</tr>
<tr>
<td>Requirement 32</td>
<td>Agents must be identified by name and biographical information.</td>
</tr>
<tr>
<td>Requirement 33</td>
<td>Agents must be identified by a unique identifier.</td>
</tr>
<tr>
<td>Requirement 34</td>
<td>Agent descriptions should interlink with resources related to the agent.</td>
</tr>
</tbody>
</table>

- Use Cases and Requirements will be published at [http://asch.wiki.gwdg.de/index.php/Main_Page](http://asch.wiki.gwdg.de/index.php/Main_Page) at the end of September 2015
Entities and Definitions (work in progress)

- Metadata Set = a description of a resource
- Item = a thing in a collection described by metadata
- Evidence = a resource proving the reliability of a statement about a resource.
- Event = an activity in the lifecycle of a resource
- Time = a time-span related to a resource via an activity or as a topic
- Agent = a person, organisation or group related to a resource via an activity or as a topic
Entities and Definitions (work in progress)

• Place = a geographic location related to a resource via an activity or as a topic

• Digital Representation = a digital image or other digital resource depicting an item

• Collection = an aggregation of items

• Statement = a predication about an item

• Holding = a description of an agents inventory information and/or access information for an item (definition by Carsten Klee, see https://wiki.dnb.de/display/DINIAGKIM/Definition+of+Holding)

• Concept = a term from a vocabulary used as a value in an item description
Information About Datasets and Access
Information about Types and Subjects
Provenance of Items
Provenance of Statements

OR
Next Steps

• Go on with the framework

• Publishing the digitized evidences

• Transcription of the evidences using TEI

• Extending or developing Application Profiles for the description of the ASCH-Entities

• Testing of the modell and profiles using different software
  – WissKI
  – Ontowiki
  – …
Any suggestions or questions?

Contact: Stefanie Rühle (mailto: sruehle@sub.uni-goettingen.de)