

Data Processing at AMICO

**AMICO Member Contribution
Management System
Brad Dietrich**

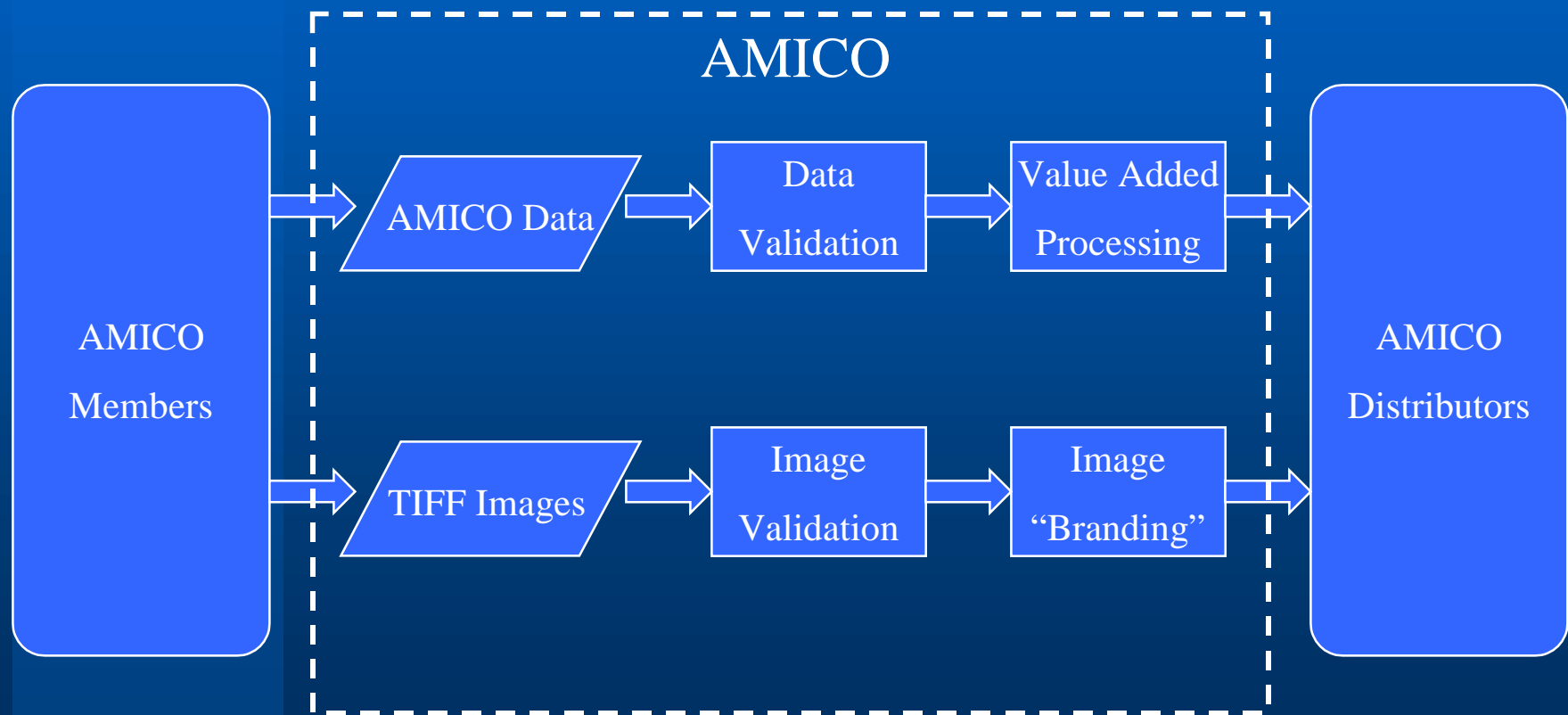
Project Goals

- **Provide a system capable of managing all aspects of the AMICO data**
 - Submission of data
 - Withdrawal of works
 - Editing/Normalization of data

AMICO Data

- **1998 Library**
 - 19,000 unique works
 - 23 institutions
- **1999 Library**
 - 55,000 unique works
 - 26 institutions
 - 200 GB of data and images

AMICO Data Flow



AMICO Process Benefits

- **Data Validation**
- **Value Added Processing**
 - Individual record editing
 - Parsing of free text fields
 - Data normalization
 - Data enrichment
- **Image Handling**
- **Distribution Management**

Data Validation

- Autonomous Validation

AMICO File Validation	
View Files Upload New Files via Browser Upload New Files via FTP Validate Text Files Validate Images Display Media Log	File Submission Menu Please select an option from the menu at the left.

Value Added Processing

- Individual Record Editing
- Parsing of Free Text Fields
- Data Normalization
- Data Enrichment

Individual Record Editing

AMICO Record Editing					
Record Editing Withdraw Record File Submission Record Linking Data Parsing Report Term Occurrences GII Matching	Find Records: <table border="1"><thead><tr><th>Catalog Records</th><th>MetaData Records</th></tr></thead><tbody><tr><td>AID: <input type="text"/> <input type="checkbox"/> Only display invalid Records <input type="button" value="Submit"/> Create a new Catalog Record</td><td>XID: <input type="text"/> <input type="checkbox"/> Only display invalid Records <input type="button" value="Submit"/> Create a new MetaData Record</td></tr></tbody></table>	Catalog Records	MetaData Records	AID: <input type="text"/> <input type="checkbox"/> Only display invalid Records <input type="button" value="Submit"/> Create a new Catalog Record	XID: <input type="text"/> <input type="checkbox"/> Only display invalid Records <input type="button" value="Submit"/> Create a new MetaData Record
Catalog Records	MetaData Records				
AID: <input type="text"/> <input type="checkbox"/> Only display invalid Records <input type="button" value="Submit"/> Create a new Catalog Record	XID: <input type="text"/> <input type="checkbox"/> Only display invalid Records <input type="button" value="Submit"/> Create a new MetaData Record				

Parsing of free text fields

AID	OCT	Results
JPGM.54.PA.1	Mid 1640s	('1643', '1646')
JPGM.55.AB.12	480 B.C.	('0480', '-0480')
JPGM.55.DA.2	ca. 1750	('1740', '1760', 'circa')
JPGM.55.DA.4	ca. 1770 - 1775	WARNING: Ignoring circa for date range! ('1770', '1775', 'circa')
JPGM.55.DA.5	1769	('1769', '1769')
JPGM.57.AA.19	Mid 1st century A.D.	('0033', '0066')
JPGM.57.PA.3	ca. 1780	('1770', '1790', 'circa')
JPGM.58.AA.1	90 A.D.	('0090', '0090')
JPGM.67.DA.10	ca. 1725	('1723', '1727', 'circa')
JPGM.67.DA.9	ca. 1785 - 1788	WARNING: Ignoring circa for date range! ('1785', '1788', 'circa')
JPGM.68.PB.4	ca. 1496	('1494', '1498', 'circa')
JPGM.69.PA.12	ca. 1660	('1650', '1670', 'circa')
JPGM.69.PA.13	1650	('1650', '1650')

Data Normalization

Occurrence Counts for Values of OTY

Field Value	Occurrence Count	Institutions Using
?	1	AGO_
Alabastra	1	JPGM
Albumen	14	JPGM
Ambrotypes	1	JPGM
Amphorai	5	JPGM
Amulet or Scarab	11	FASF
Amulets	20	CMA_, MIA_
Antefixes	2	JPGM
Appliqué	1	JPGM
Architectural	9	FASF
Architectural Fragment	1	AIC_
Architecture	1	PMA_
Armchairs	2	JPGM
Armor	2	JPGM
Arms and Armor	146	CMA_
Arms or Armor	1	FASF
Aryballoi	3	JPGM
Askoi	1	JPGM
Assemblage (objet trouvé)	1	MACM
Assemblage photographique, électrique	1	MACM
Assemblage pictural	1	MACM
Assemblage sculptural, photographique	1	MACM

Data Enriching

- **Matching to Artist Authorities**
- **Extending Artist Authorities / ARS**
- **AAT & TGN Plans**
 - Using Hierarchy
- **Enriching Searches Through Authorities**

Image Handling

- Image Validation
- Image “Branding”
 - TIFF headers
 - Copyright, owner, artist, and work information
- Thumbnail Creation
- External Metadata Creation
- Record Matching

Distribution Management

- **Keeping Distributors Library Current**
 - Weekly Updates
- **Tracking Changes**
- **User Permissions**
- **Public Thumbnail Web Site**

The Future

- **Editorial Decisions about Normalization of Terms**
- **Artist Biography Authority**
- **Parsing of Other Free Text Fields**
- **Publish XML/RDF**