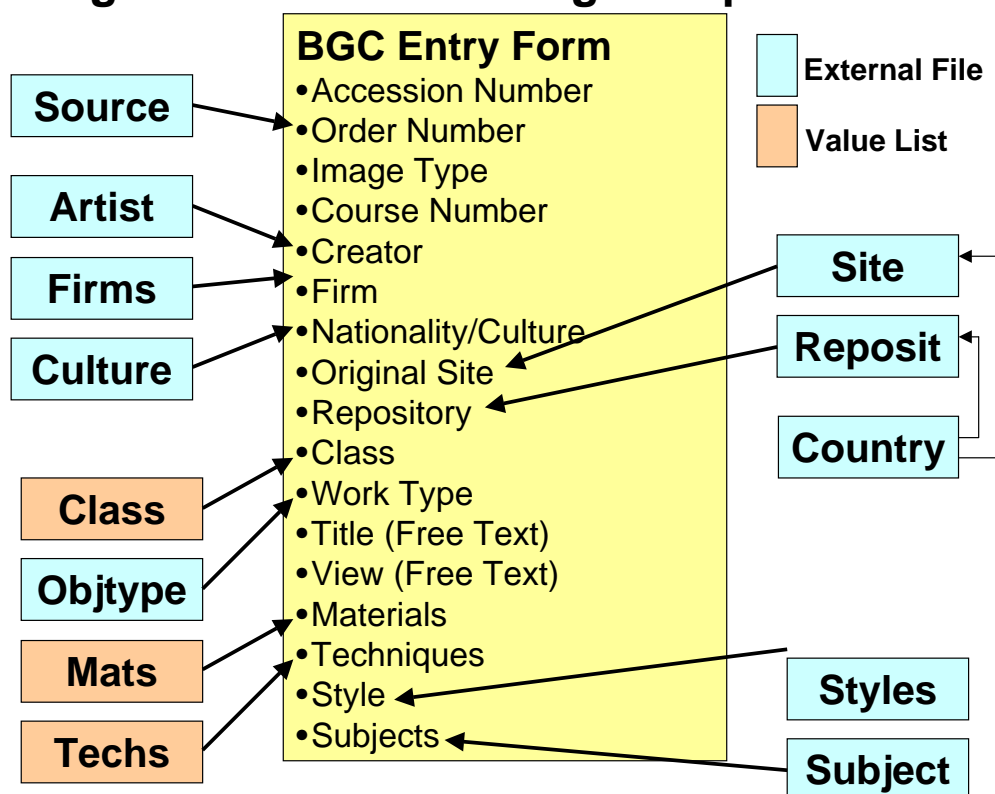


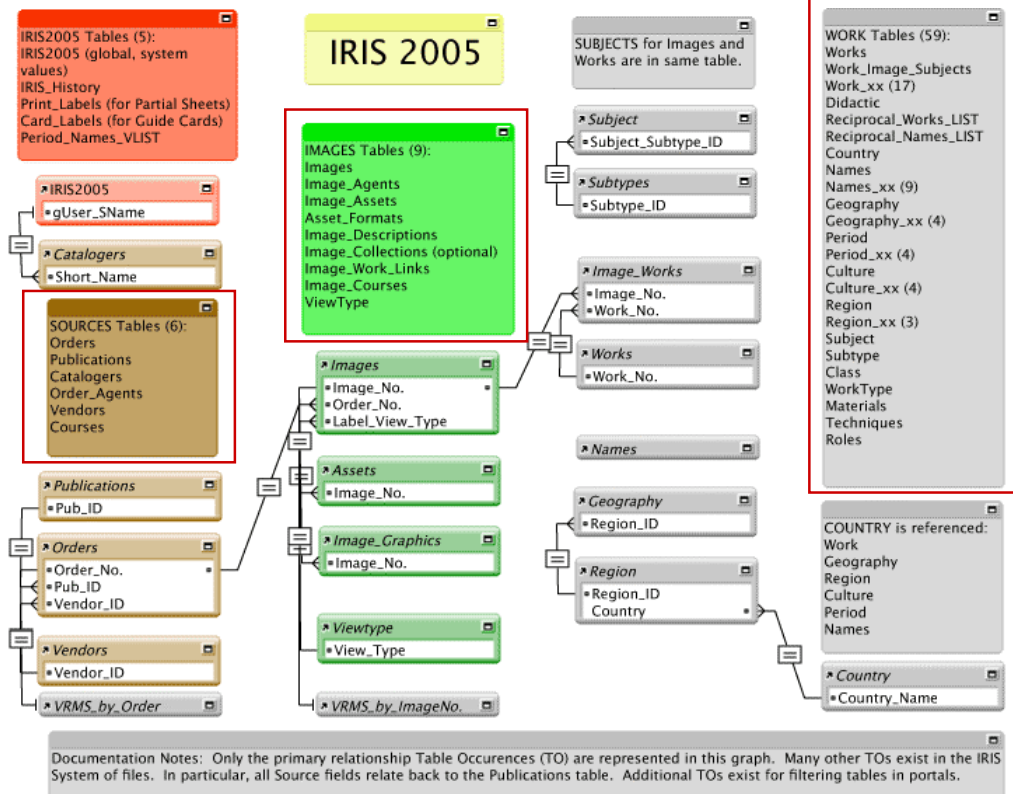
From Flat to Relational: Data Migration Strategies for a Small Collection

Johanna Bauman
Bard Graduate Center
Kansas City, March 28, 2007

Existing Data Structure Using Multiple Linked Files



IRIS2005 Relational Table Structure (Simplified)



State of the BGC Databases Before Migration

- c. 40,000 records in the “Yale (a.k.a. Susan) Template” on the FileMaker 6.0 platform (25,000 of these imported from a defunct VRMS database with only Access files)
- c. 2,000 image and 1,600 work records in old version of IRIS that used linked files rather than a relational structure on the FileMaker 6.0 platform

Steps for Data Migration

- I. Planned and Developed Strategy
- II. Mapped Data
- III. Cleaned Up Data
- IV. Marked Work and Image Records
- V. Imported Files into Individual Tables and Created Relationships
- VI. Initial Data Clean Up Using Scripts
- VII. Developed Plan for Ongoing Clean Up
- VIII. Planning for the Next Migration....

I. Planned and Developed Strategy (Goals & Assumptions)

Goals

- Bard and BGC would have authority records that were consistent with one another
- Import records directly into the main image/work structure in IRIS rather than putting the legacy data into an intermediary place

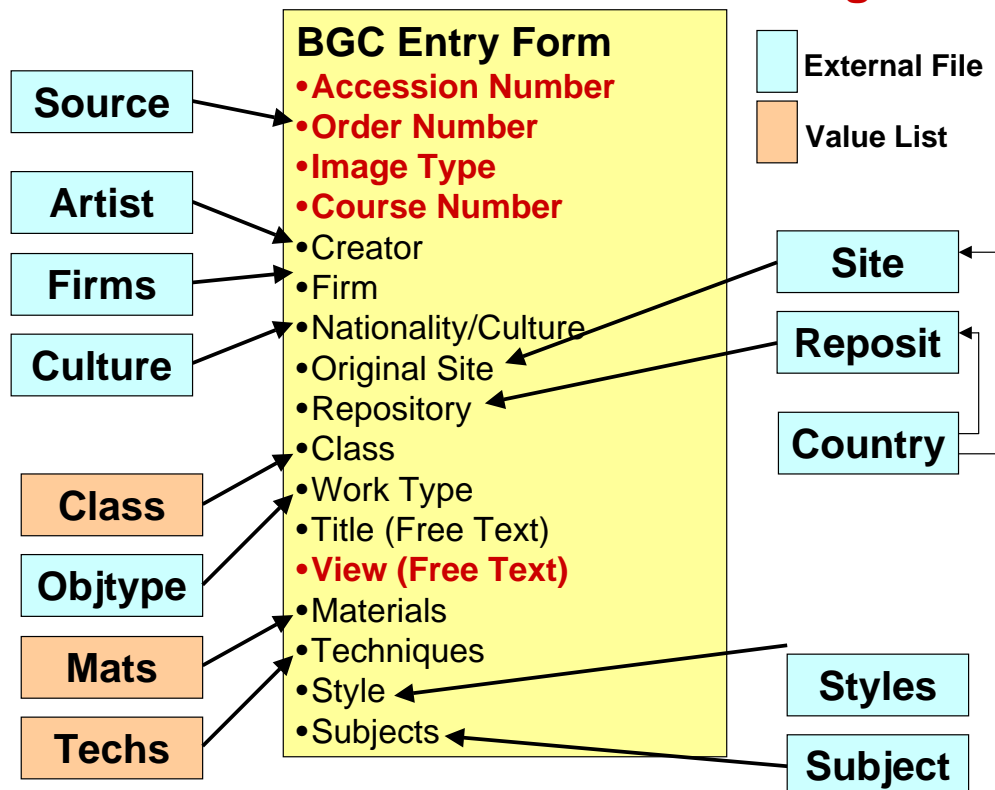
Assumptions

- All records would have to be cleaned up before being uploaded into MDID as part of the process of going digital
- Some data is better than no data
- Preserving data structure was paramount

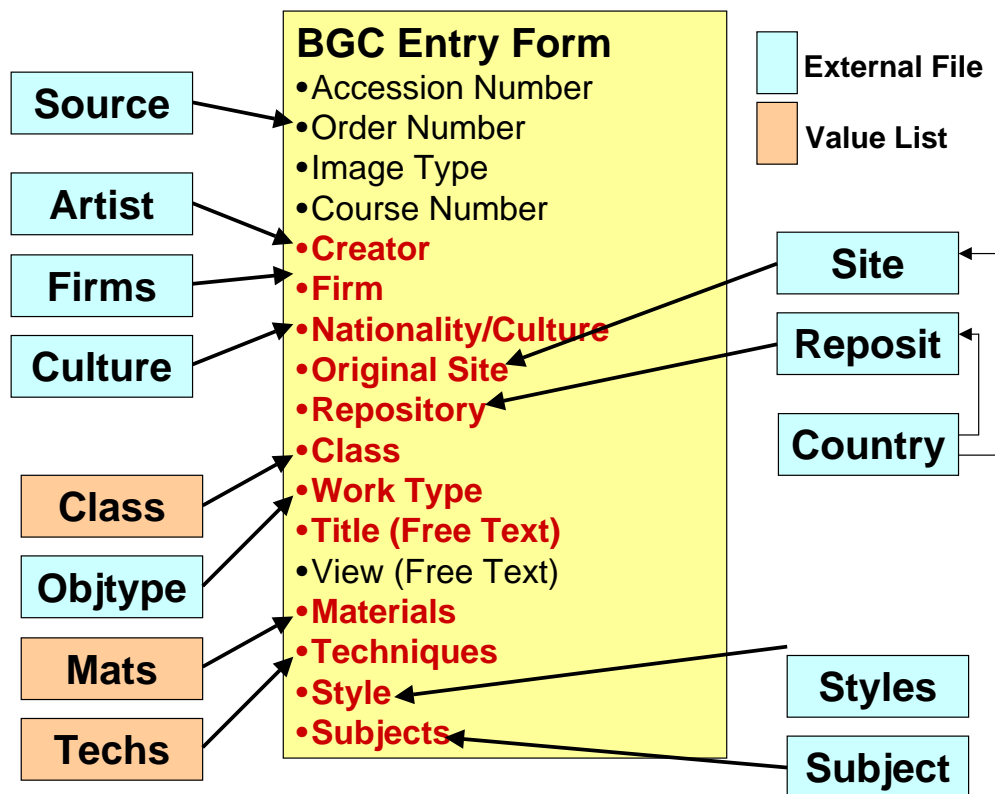
I. Planned and Developed Strategy (Steps)

- Familiarized myself with the VRA Core
- Looked at data and identified what fields in use correspond to the VRA Core fields
- Analyzed the database and broke it down to image vs. work
- Thought about what information was repeated across fields and what was unique
- Made arrangements to work with a consultant
- Learned as much as possible about how FileMaker works and developed searching strategies

Identified Fields Associated with the **Image** Record



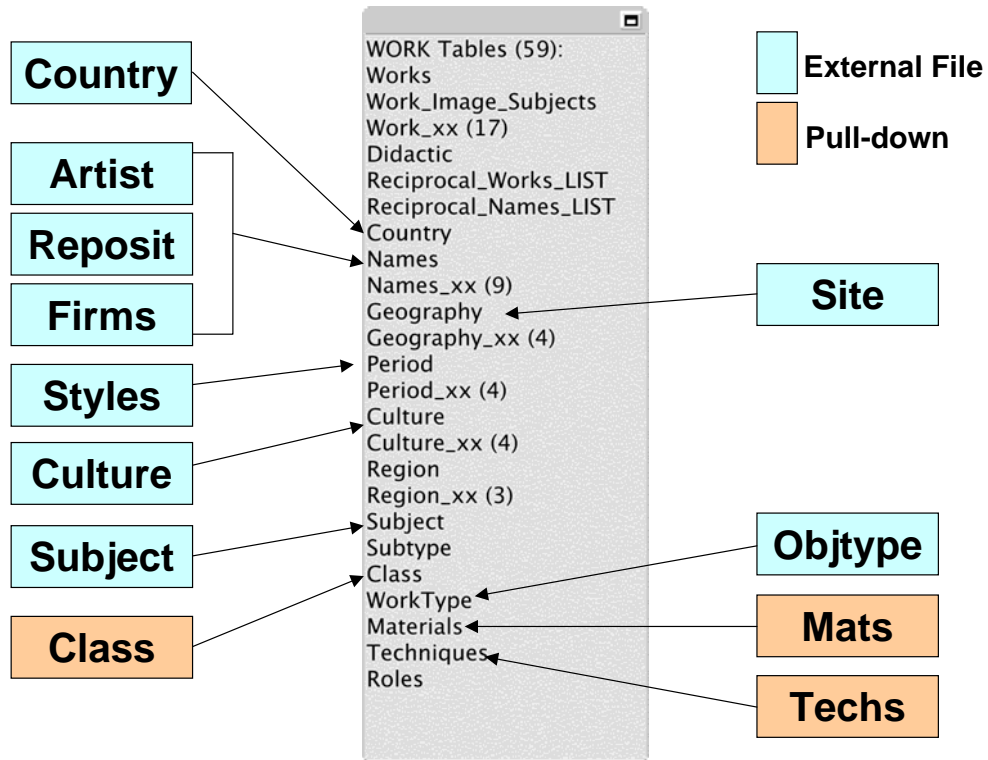
Identified Fields Associated with the **Work Record**



II. Mapped Fields

- Analyzed the new data structure and the old data structure
- Determined what fields/pull downs/files corresponded to authority tables in new system
- Created maps for all files and tables

Identified Tables Associated with the **Work Record**



II. Mapped Authorities

Sample Mapping for Artist Authority Table

IRIS: Names Table	BGC: Artist File
Periods	Artist/Era Culture
Names	Artist Sort Name
Cultures	Associated Nationality
Vocab	Authority
combined into dates	B Descriptor
combined into dates	BCE
Start	Birth Date
Countries	Country
combined into dates	D Descriptor
Entry Date	Date Modified
End	Death Date
Gender	Gender
Cataloger	Modified By
Notes	Notes

III. Cleaned Up Existing Data

- Identified duplicate source records and consolidated images connected to them
- Tried to make sure all creators and repositories had authority records even if they were incomplete
- Exported records and cleaned them up using Excel to standardize terms and names
- Filemaker Pro is very flexible and allows you to easily add fields to assist in the migration

IV. Marked Work and Image Records


- Added a field to mark work records and a field for entering a work record number
- Developed a strategy for distinguishing between work and image records
- Spent c. 3-4 weeks of staff time on the project
- At the end of that time whatever records had not been marked as work records became image records by default
- Since the Site file contained a lot of architectural “works” these were also marked to distinguish between architectural works and true geographic locations

Records Grouped by Creator in Table View

FileMaker Pro - [Entry.fp5]

File Edit View Insert Format Records Scripts Window Help

Entry F...	Local Class 1	Cataloger	Page Number	Creator Sort Name	Label Title	Creat...	Order Nu...	Repository Long ...
	Painting	DMV		Raphael	An Allegory	painter	371	London, National
	Painting			Raphael	An Allegory	painter	771	London, National
	Painting	JRS		Raphael	cartoon for An	painter	771	London, National
	Painting	AAR		Raphael	Cartoon of Christ	painter	2243	London, Victoria
	Painting	AAR		Raphael	Cartoon of St.	painter	2243	London, Victoria
	Painting	AAR		Raphael	Cartoon of the	painter	2243	London, Victoria
	Painting	AAR	31	Raphael	Christ's Charge to	painter	1351	London, Victoria
	Drawing			Raphael	Christ's Charge to	painter	1351	
	Textile	AAR	38	Raphael	Christ's Charge to	painter	1351	Vatican City,
	Textile	AAR	211	Raphael	Conversion of	painter	1347	Vatican City,
	Textile	AAR	213	Raphael	Conversion of	painter	1347	Vatican City,
	Painting	AAR	34	Raphael	Conversion of			
	Textile	AAR	43	Raphael	Conversion of			
	Textile	AAR	44	Raphael	Death of Anan			
	Painting	AAR	33	Raphael	Death of Anan			
	Drawing	AAR	pL 8	Raphael	design for a bo			
	Drawing	AAR	pL 9	Raphael	design for a bo			



Marked Work Record

FileMaker Pro - [Entry.fp5]

File Edit View Insert Format Records Scripts Window Help

Entry Fo... Flag Proofed Missing Duplicate Old VRMS Information Export

Go First Go Next Go Prev Go Last View List Images Go Source Dupe or New Image Find Record Print Record Print This Label 127 Records Work Record Number New Work Record Number

Records: 50275 Found: 127 Sorted

Order, Record, and File Information

Order Number	Creation Date	Last Updated	Cataloger	Accession Number	Source Number/Name
371	1/9/2002	11/17/2005	DMV	7440	

Image AWS File Name: bpg7440.tif File Name: CD Number: Vendor Number: Work Record

Image Format: slide, 35 mm scanned slide, tif digital original, tif duped slide, 35mm scanned slide, jpg digital original, jpg

Makers and Associated Persons

Creator Sort Name: Raphael Birth and Death Dates: 1483-1520 Creator Role: painter Associated Nationality: Italian Attribution: Probability: Additional Creator: Ad: Creator Attribution: Ad: Creator Probability: All Creators (Search Field): Raphael; Ad: Associated Firm: Firm Start and End Dates: Nationality Culture: Nationality Culture Dates:

Drawer and Label Information

Label Type: Known Artist Local Class 1: Painting Local Class 2: Italy Local Class 3: 15th-16th Local Class 4: PAINTING Raphael 15th-16th 1483-1520 ITALY An Allegory (Vision of a Knight) c.1504; London, National Gallery, -- 7440 Label Title: An Allegory (Vision of a Knight)

Entered Numbers for Image Records

FileMaker Pro - [Entry.fps]

File Edit View Insert Format Records Scripts Window Help

Entry Fo... Flag Proofed **Image Data Entry** Missing Duplicate **Old VRMS Information** Export

Records: 50275 Found: 127 Sorted

Go First Go Next Go Prev Go Last View List Images Go Source Duplicate or New Image Find Record Print Record Print This Label Print 117 Labels EXPORT 117 Records

Order, Record, and File Information Work Record Number: 7440

Order Number	Creation Date	Last Update	Accession Number	Course Number/Name
771	4/28/2003	11/17/2005	13924	

Image AIS File Name: bpe13024.Bf File Name: CD Number: Vendor Number: Image Format: slide, 35 mm scanned slide, tif digital original, tif duped slide, 35mm scanned slide, jpg digital original, jpg Page: 63 Plate: Figures:

Makers and Associated Persons Work Record

Add Creator Sort Name: Raphael Birth and Death Dates: 1483-1520 Creator Role: painter Associated Nationality: Italian Attribution: Probability: Additional Creator: Ad: Creator Attribution: Ad: All Creators (Search Field): Raphael; Ad: Associated Firm: Firm Start and End Dates: Nationality Culture: Nationality Culture Dates:

Old VRMS Information

VRMS Number: 1138-93 Accession Number: 1138 Order Number: 58 VRMS Dates: VRMS country: Italy Requestor: Frank, Isabelle Author: Dunkerton, Jill, et al Title: Giotto to Durer Page Number: 371

Drawer and Label Information

Label Type: Known Artist Local Class 1: Painting Local Class 2: 16th Local Class 3: Italy Local Class 4: PAINTING Raphael 16th 1483- 1520 ITALY An Allegory (Vision of a Knight) c.1504; London, National Gallery, -- 13924 Label Title: An Allegory (Vision of a Knight) tannars

Split Titles of Image Records to Create View Fields

Object Description

Classification Term: Ap. Arts: Furniture

Narrow Term (Search Field):

Object Type: furniture: support furniture: tables, console tables

Add

Title: Louis XV console table, detail of putti

View:

ViewType:

Description:

Century: 18th

Reimported Split Title to View and View Type Fields

Object Description

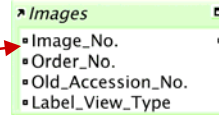
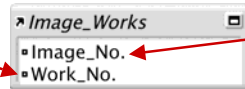
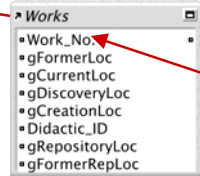
Classification Term	Ap. Arts: Furniture			
Narrow Term (Search Field)				
Object Type	furniture: support furniture: tables, console tables			
	Add			
Title	Louis XV console table			
View	detail of putti			
ViewType	detail			
Description				
Century		18th		
Object Date	c.	1730		

V. Imported Files into New System

- Brought on a consultant to work on scripting
- Consultant had worksheets and checklists that went through step by step for us to follow

Relationship Between Work and Image in IRIS2005 Using Relation Tables

Works File



Images File



V. Imported Files into New System

Import Order

- Authorities
- Orders
- Publications/Sources
- Works
- Images

List of Tables to be Imported

INSTALL 2005

Install Menu to import data to Files & Authority Tables

Authorities
IRIS 2005

Main Menu >> IRIS History
Review Install Layout in IRIS2005

Order Agents
Value Lists in IRIS2005
Value Lists in Sources

Authorit >> Publications

Countries

Periods
Period - Value List in IRIS2005

Cultures
Periods-return for Assoc. Cultures

View Types

Vendors

Work Types

Materials

Techniques

Subtypes

Subjects

Didactic Labels

Roles (was Creator Value List)

Classes (was Categories)

Regions

Site >> Geography

Orders (Source)

Publications (Titles)

Creator >> Names
(then Groups & Members)

Repositories

VRMS (can be done at any time...import all records, set by matching field names)

Works

All Tables

WSLinks

Images (Surrogat)

Digital to Assets
Table View

Courses (own text file)

Importing Fields to the "Work Record"

FileMaker Pro - [INSTALL2005]

Summary **All Related Tables for Work Record** ly & gID used for Duplicate script & make related sbscript **INSTALL Menu**

Work No. 524909 Label Title Seated Figure Original_Work_No. 37809 gID - new work # **List w/Did.**

Classes Work_No. 139 524909 **Classes** Title Seated 524909 **Titles**

Related **Import WRelated** 1st Find All

WorkTypes Work_No. **WorkTypes** Texts **Import Texts** Find 1st

Work_Agents 29638 524909 **Agents** Descriptions **Descriptions & Notes** Find 1st

Dates Type 1917 524909 Creation **Dates**

Centuries 20th C. 524909 **Centuries** Find 1st

Subjects **Subjects**

Countries **Countries** Find 1st

Cultures **Cultures** Find 1st

Periods **Periods** Find 1st

Techniques **Techniques** Find 1st

Materials 1204 524909 **Materials** Find 1st

Locations 2176 524909 **Locations** Built in multiple passes

Measurements **Measurements** New - will be built

References **References** New - will be built

(can take a while, may want to run "at lunch")

VI. Initial Data Clean Up Using Scripts

Table Showing Unlinked Creators

The screenshot shows the FileMaker Pro interface with a table titled 'wCreator to Work Agents'. The table has the following columns: ID, Work No., Names_ID, Attribution, temp, Probability, :-Name, CreatorName_imported, and Agent_wAttribution. The data shows a list of creators with their respective IDs and work numbers, all of whom are currently unlinked (empty cells in the Name and Attribution columns).

ID	Work No.	Names_ID	Attribution	temp	Probability	:-Name	CreatorName_imported	Agent_wAttribution
25962	500016	29640				Donatello	Donatello	Donatello
25963	500017	29640				Donatello	Donatello	Donatello
28915	502988	29648				Donatello	Donatello	Donatello
33819	508270	29648				Donatello	Donatello	Donatello
39988	514516	29640				Donatello	Donatello	Donatello
39993	514522	29640				Donatello	Donatello	Donatello
42533	517326	29648				Donatello	Donatello	Donatello
55388	532592	29648				Donatello	Donatello	Donatello
55389	532490	29648				Donatello	Donatello	Donatello
55440	532490	29640				Donatello	Donatello	Donatello
55456	532490	29648				Donatello	Donatello	Donatello
55457	532490	29648				Donatello	Donatello	Donatello
55458	532490	29648				Donatello	Donatello	Donatello
55459	532490	29640				Donatello	Donatello	Donatello
55800	532898	29648				Donatello	Donatello	Donatello
55881	532898	29648				Donatello	Donatello	Donatello
55882	532898	29648				Donatello	Donatello	Donatello
55924	532592	29640				Donatello	Donatello	Donatello
55925	532490	29640				Donatello	Donatello	Donatello
60037	532708	29648				Donatello	Donatello	Donatello
58805	532592	29648				Donatello	Donatello	Donatello
58806	532592	29640				Donatello	Donatello	Donatello
58807	532592	29640				Donatello	Donatello	Donatello
58808	532592	29648				Donatello	Donatello	Donatello
58809	532592	29648				Donatello	Donatello	Donatello
58810	532592	29648				Donatello	Donatello	Donatello
58811	532592	29640				Donatello	Donatello	Donatello
59608	534192	29648				Donatello	Donatello	Donatello
60503	532709	29648				Donatello	Donatello	Donatello
60901	537015						Donatello	
60902	537016						Donatello	
62983	537041						Donatello	
62984	537042						Donatello	
63026	537086						Donatello	
63027	537086						Donatello	
63028	537087						Donatello	
63029	537088						Donatello	
63275	537357						Donatello	
63276	537358						Donatello	
63277	537359						Donatello	

Same Table after Running the Update Script

The screenshot shows the same FileMaker Pro interface after running an update script. The table 'wCreator to Work Agents' now has data populated in the previously empty 'Attribution', 'temp', and 'Probability' columns. The ':-Name' column now contains the names of the creators, and the 'Agent_wAttribution' column is also populated.

ID	Work No.	Names_ID	Attribution	temp	Probability	:-Name	CreatorName_imported	Agent_wAttribution
25962	500016	29640				Donatello	Donatello	Donatello
25963	500017	29640				Donatello	Donatello	Donatello
28915	502988	29648				Donatello	Donatello	Donatello
33819	508270	29648				Donatello	Donatello	Donatello
39988	514516	29640				Donatello	Donatello	Donatello
39993	514522	29640				Donatello	Donatello	Donatello
42533	517326	29640				Donatello	Donatello	Donatello
55388	532592	29648				Donatello	Donatello	Donatello
55389	532490	29648				Donatello	Donatello	Donatello
55440	532490	29648				Donatello	Donatello	Donatello
55456	532490	29640				Donatello	Donatello	Donatello
55457	532490	29648				Donatello	Donatello	Donatello
55458	532490	29648				Donatello	Donatello	Donatello
55459	532490	29648				Donatello	Donatello	Donatello
55800	532898	29648				Donatello	Donatello	Donatello
55881	532898	29648				Donatello	Donatello	Donatello
55882	532898	29648				Donatello	Donatello	Donatello
55924	532592	29648				Donatello	Donatello	Donatello
55925	532490	29640				Donatello	Donatello	Donatello
55937	532708	29640				Donatello	Donatello	Donatello
58805	532592	29648				Donatello	Donatello	Donatello
58806	532592	29648				Donatello	Donatello	Donatello
58807	532592	29648				Donatello	Donatello	Donatello
58808	532592	29640				Donatello	Donatello	Donatello
58809	532592	29648				Donatello	Donatello	Donatello
58810	532592	29648				Donatello	Donatello	Donatello
58811	532592	29648				Donatello	Donatello	Donatello
59608	534192	29648				Donatello	Donatello	Donatello
60503	532709	29640				Donatello	Donatello	Donatello
60901	537015	29648				Donatello	Donatello	Donatello
60902	537016	29648				Donatello	Donatello	Donatello
62983	537041	29640				Donatello	Donatello	Donatello
62984	537042	29640				Donatello	Donatello	Donatello
63026	537086	29648				Donatello	Donatello	Donatello
63027	537086	29648				Donatello	Donatello	Donatello
63028	537087	29648				Donatello	Donatello	Donatello
63029	537088	29648				Donatello	Donatello	Donatello
63275	537357	29648				Donatello	Donatello	Donatello
63276	537358	29648				Donatello	Donatello	Donatello
63277	537359	29648				Donatello	Donatello	Donatello
63310	537392	29648				Donatello	Donatello	Donatello
63311	537393	29640				Donatello	Donatello	Donatello

VII. Developed Plan for Ongoing Clean Up

- Clean up records on an as-needed basis with data dumped into import and note fields
- Create authorities for unlinked creators
- Fill in authorities with complete information and sources

Layout Showing Imported, Unlinked Data for Creator

The screenshot shows the FileMaker Pro interface for the IRIS 2005 database. The layout is titled "Work Agents" and is part of the "Bard Graduate Center" database. The main record is for "Work No. 218303" with the title "armchair for Robert Child(from a suite made after Adam design) attr:J.Linnell marquetry with".

The layout is divided into several sections:

- CULTURES:** A table with columns for ID, Culture Name, Culture Context, and Label. The first row shows ID 1049, Culture Name "English", Culture Context "European", and Label "X". Below this are two "Import_CultureName" fields, one containing "English".
- AGENTS:** A table with columns for ID, Agent Name, Attribution, Probability, and Gender. The first row shows ID 38094, Agent Name "Adam, Robert", Attribution "Filter Agents", Probability, and Gender "male". Below this are three rows of "Work Role", "Extent", and "Culture" fields. The first row shows "116", "artist", and "1053 British". The "CCO Agent Display" field shows "Artist: Robert Adam (British, 1728 -1792)". The "CreatorName_Imported" field shows "Adam, Robert".

The layout also includes a "Loop" button and a "Notes" field.

VII. Planning for the Next Migration....

- Keep up to date on changes in VRA Core and with the system you are using
- Be consistent and keep cataloging clean to make next migration easier
- Learn from your mistakes

What we learned...

- Migration will always take longer than you think - Initially projected 2 weeks, it actually took more like two months
- You will have to stop cataloging for a time, need to come up with alternate projects for staff and student workers
- If you do have to go back to cataloging in the old system before the migration is over, be sure to catalog in a way that conforms to the standard to which you are migrating