

Proposal for Conversion of the
Bancroft Library Picture Collection Catalog
to Machine Readable Form

Joan S. Mitchell
Michael D. Cooper

School of Librarianship
Univeristy of California
Berkeley, California

October 1974

Preface

There are several possible courses of action that can be taken to improve the Bancroft Library picture catalog. This paper analyzes one alternative - conversion of the catalog to machine readable form so that a book catalog can then be produced. Before this alternative or any others is adopted, it is suggested that a thorough cost analysis of all alternatives be conducted.

PART I

Description of the Existing File

Introduction

The Bancroft Library of the University of California, Berkeley contains an excellent collection of materials related to Western America. One of its major assets is an extensive collection of photographs, which are accessed by a subject catalog. The catalog is in the form of approximately 17,700 cards contained in 14 file drawers.¹

There are a number of problems in using and maintaining the file and the collection of photographs and negatives. First, access to the collection is only by subject. There is no shelf list ordered by the accession number of the photograph, nor is there any record of which subject headings have been assigned to each item. Second, through many years of use, a certain lack of uniformity has emerged in the way in which information has been recorded in the file.²

The purpose of this study is to analyze the characteristics of the picture catalog, propose a method for bibliographic description of the records in the file, and suggest a methodology for converting the records to machine-readable form so that they can be processed by computer. Once the records are in machine-readable form, computer programs can be written to produce a book catalog of the holdings, ordered by accession or location number.

¹A fifteenth drawer, "Pictures in Books," has been excluded from this study.

²For example, there is considerable variation in the form of the accession or location number.

Bibliographic Description of the Records

Figure 1 presents five typical cards selected from the Picture Collection catalog. The first line of the first card is the subject heading "Adobe House, 205 Blithedale, Mill Valley" and "12991-12993" represents the numbers by which the photograph may be located in the picture files in Bancroft Library stacks. Similarly, on the second card the subject heading is "Carpenter, William Diarist" and under that heading there are five photographs each with an additional description.

A random sample of 20 cards in each of 12 drawers was taken in order to determine the elements that comprise each record in the catalog. Eleven fields have been isolated as being representative of the elements forming each record.³ The fields are listed in Table 1 with their frequency of occurrence within the sample. The same five sample cards appear in Figure 2 with the different fields illustrated. A discussion of each field follows.

Subject Heading, Description (1), Description (2)

Subject heading denotes the first line (or sentence) on the card. In the present catalog, the cards are arranged alphabetically by subject heading. Description (1) and (2) are arbitrary names given to the one or two lines of further refinement of the subject. In all but 3 instances in the sample the subject heading and descriptions were applicable to all location numbers found on the card.

A frequency distribution of the number of characters in the Subject Heading and Descriptions (1) and (2) can be found in Figures 3, 4, and 5 respectively.

³Occasionally a bracketed number appears in the lower right hand corner of a card. This is an old accession number and is to be ignored.

Figure 1 Part 1
Sample Catalog Cards

Adobe House 205 Blithedale , Mill Valley
three exterior photos picture drawer
12991-12993

gift of Mrs. Lucretia H. Little 5.25.62
enclosed in envelope are notes suggesting
its origin.

Carpenter, William Diarist
his sketches of places in Nicaragua and Mexico
photonegatives picture drawer
Monkey Point 17110(141)neg
Lake Nicaragua and Mts. 17110(142)neg
San Juan del Sud harbor 17110(143)neg
Cape Horn 17110(144)neg
Acapulco Harbor 17110(145)neg

all from C-B812 qv

Carpinteria, California - Views
Giant Grape Vine trunk
photo 11368 picture drawer

Figure 1 Part 2
Sample Catalog Cards

Crater Lake, Oregon
see Gaytes Collection



Edouart, A. Artist
Headquarters of the Mendocino Indian Reservation
Original
Honeyman 63.2.228C

Picture drawer
16903(402)4x5 color



Table 1
 Frequency of Occurrence of Fields in
 Bancroft Library Subject Picture Catalog Sample

Field Name	Frequency of Occurrence
Subject Heading	240
Description (1)	161
Description (2)	20
Type of Holding	153
Location Number	229 ⁴
Location of Holding	214
Collection	29
Collection Number	22
Note	9
Gift	8
See Reference	9

⁴Two cards gave the title and issue of a magazine instead of a location number, and 9 cards were "see references" without location numbers.

Figure 2 Part 1

Sample Catalog Cards With Fields Annotated

Subject Heading → Adobe House 205 Blithedale , Mill Valley
 Description (1) → three exterior photos picture drawer ← Location of Holding
 12991-12993 ← Location Number
 Gift → gift of Mrs. Lucretia H. Little 5.25.62
 Note → { enclosed in envelope are notes suggesting
 its origin.

Subject Heading → Carpenter, William Diarist
 Description (1) → his sketches of places in Nicaragua and Mexico
 Type of Holding → photonegatives picture drawer ← Location of Holding
 Description (2)'s → { Monkey Point 17110(141)neg
 Lake Nicaragua and Mts. 17110(142)neg
 San Juan del Sud harbor 17110(143)neg
 Cape Horn 17110(144)neg
 Acapulco Harbor 17110(145)neg } ← Location Numbers
 Collection Number → all from C-B812 qv

Subject Heading → Carpinteria, California - Views
 Description (1) → Giant Grape Vine trunk
 Type of Holding → photo 11368 picture drawer ← Location of Holding
 Location Number

Figure 2 Part 2
Sample Catalog Cards With Fields Annotated

Subject Heading → Crater Lake, Oregon
See Reference → see Gaytes Collection

Subject Heading → Edouart, A. Artist
Description (1) → Headquarters of the Mendocino Indian Reservation
Type of Holding → Original
Collection → Honeyman 63.2.228C
Collection Number → ↑
Picture drawer ← Location of Holding
16903(402)4x5 color ← Location Number

Figure 3

Frequency Distribution of Characters in Subject Headings Sampled

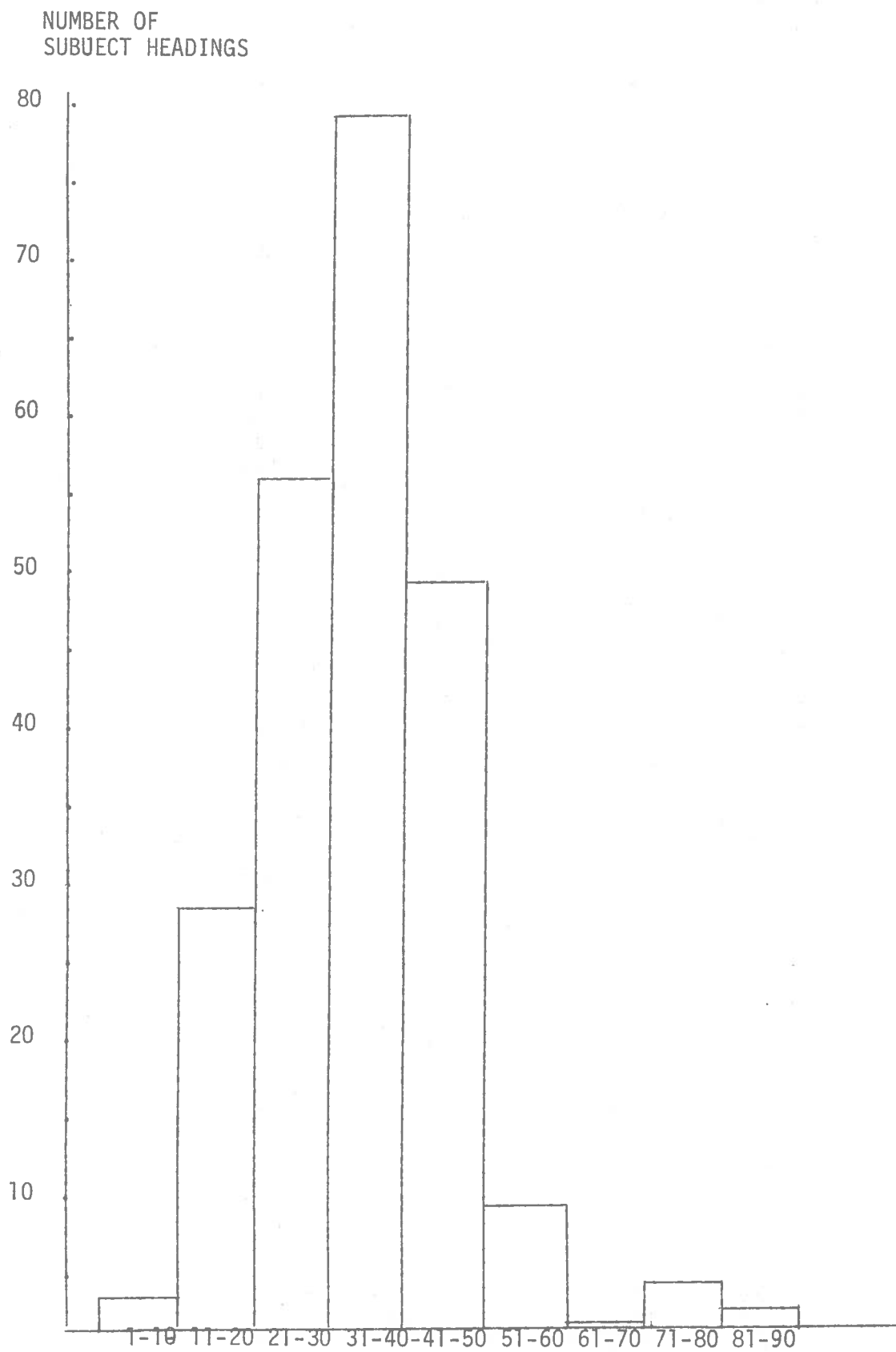


Figure 4

Frequency Distribution of Characters in
Description (1)'s Sampled

NUMBER OF
DESCRIPTION (1)'s

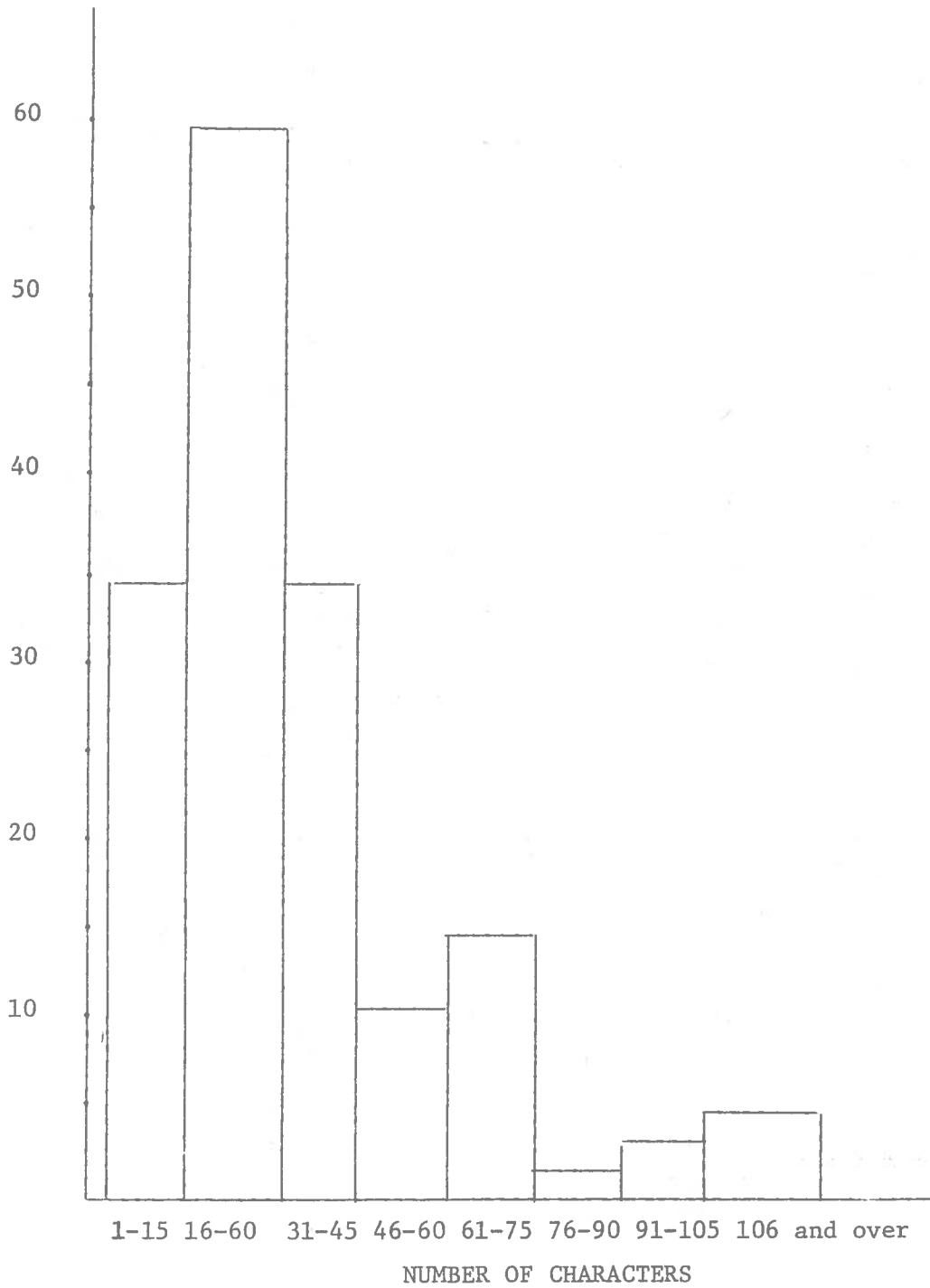
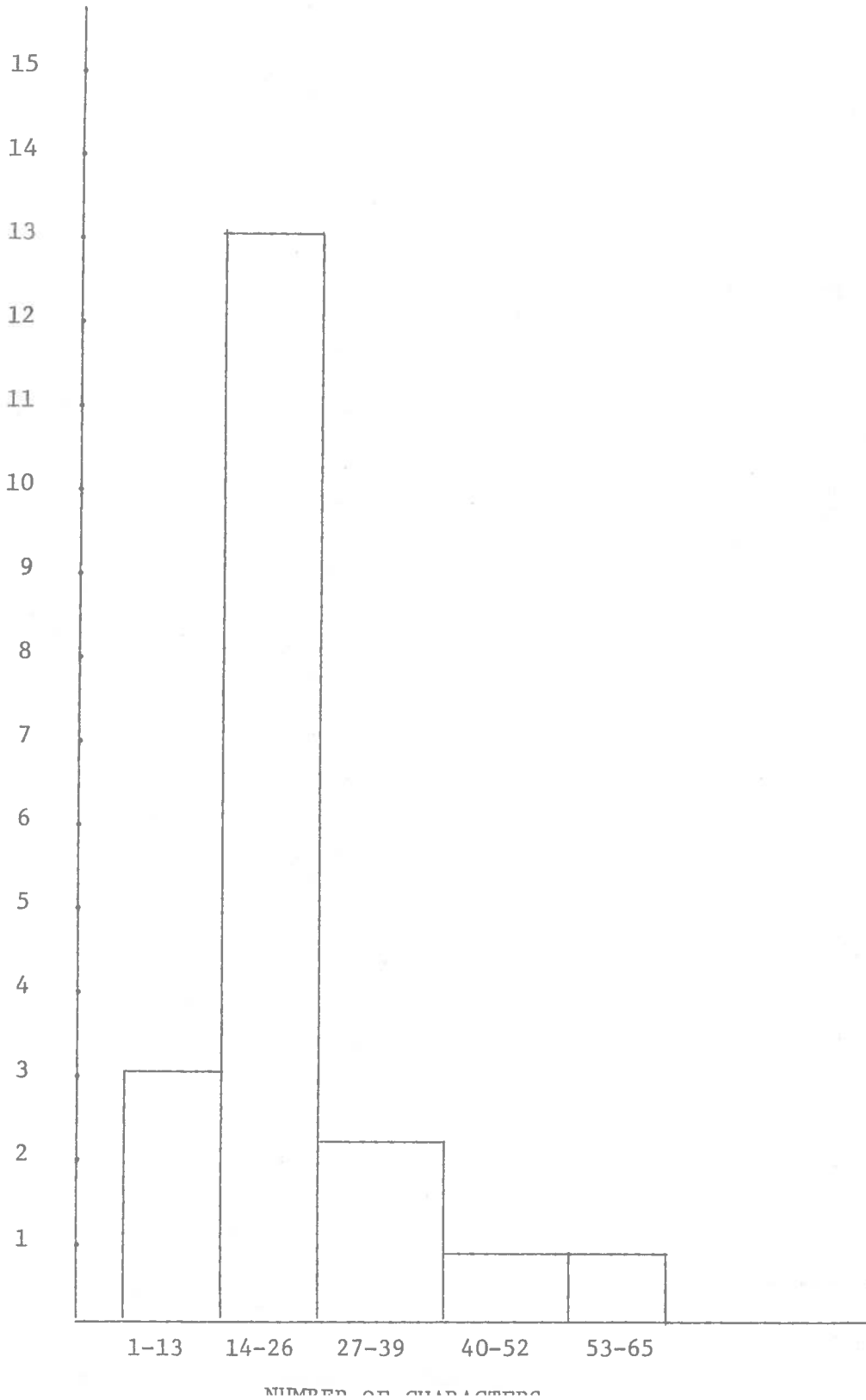


Figure 5

Frequency Distribution of Characters in
Description (2)'s Sampled

NUMBER OF
DESCRIPTION (2)'s



Type of Holding

Type of Holding is the element on the record that indicates what kind of picture is included in the collection. A wide variety of types of holding were encountered in the sample and are indicated in Table 2 with their corresponding frequency of occurrence. It should be noted that the type of holding is often included as part of the description and is not repeated as a separate element on the record. e.g. "Watkins photo", "photo of a group of numbers", etc.

Location Number

Items in the Bancroft Library Picture Collection are represented by a variety of forms of Location Numbers. Of the 229 cards with numbers in the sample, 67 had more than 1 number on it.

The various forms of numbers and their frequency of occurrence are indicated in Table 3. The frequency of occurrence is interpreted here to mean whether or not a particular form is found on a card and not how many times it occurs on a card, i.e., if a 5 digit number is listed 3 times on a card its frequency of occurrence in the sample is listed as 1.

Location of Holding

Location of Holding indicates the place where the item is physically held. The location was given on 214 of the 229 cards in the sample that had location numbers. The words and phrases used to indicate the location of the holding are given in Table 4 with their frequency of occurrence. Although listed separately, "picture" and "picture drawer" are synonymous.

Collection

The Collection field pertains to the name of the collection (e.g. Honeyman) to which the item(s) on the record belong. It is generally found in the lower left hand corner of the record.

Table 2

Frequency of Occurrence of Type of Holding in
Bancroft Library Picture Catalog Sample

Type of Holding	Frequency of Occurrence
photo	67
photos	21
photonegetive	28
photo and negative	5
stereo photo	3
print	3
photos and sketches	2
photographs	2
stereo	2
negative	2
negatives	2
negative available	2
illustration	2
litho	1
original water colors	1
color transparency	1
photos and negatives	1
photo negative (color)	1
photo, colored	1
photos, including portrait	1
portrait photo	1
postcards	1
snapshots	1
ambrotype	1
picture	1
original	1

Table 3 Part 1

Frequency of Occurrence of Various Forms of Location
Number of Bancroft Library Picture Catalog Sample

Form of Location Number	Frequency of Occurrence
XXX ⁵	11
XXXX	66
XXXXX	52
XXXX ^a	6
XXXX neg	1
XXXXX neg	8
XXXXX stereo	3
X (XX)	1
XXXX (X)	2
XXXX (XX)	6
XXXX (XXX)	4
XXXXX (XX)	20
XXXXX (XXX)	16
XXXXX (XXXX)	3
XX/XXX neg	1
XXXX (XX) neg	8
XXXX (XXX) neg	27
XXXXX (X) neg	2
XXXXX (XX) 4x5 color	1
XXXXX (XXX) 35mm color	1

⁵The character "X" in this table represents a number which can take on the values 0, 1, 2, 3, 4, 5, 6, 7, 8, 9. The characters other than "X" are actually presented in the field - they are literals.

Table 3 Part 2

Frequency of Occurrence of Various Forms of Location
Numbers in Bancroft Library Picture Catalog Sample

Form of Location Number		Frequency of Occurrence
F XXX SX SXXXx	No. X	2
F XXX SX SXXXx	No. XX	4
F XXX CXXX	XX:X:XXX	1
F XXX SXXX	X:X:XX	1
F XXX MXx	X:XX:X	1
E XXX GXXx	p.XXX neg	1
F XXX WXXx	p.XXX neg	1
ff F XXX UYX	v.2x:104	1
ff F XX JXX	v.3:17x	
F XXX MXx	vol.44 3/25/88 page 1 neg	1
ff F XXX CXXX	v.3:11x	1
F XXX WXXX	v.39:22:7 neg	1
ff F XXX.X WXXXx	No.7	1
f F XXX LX LXXX	Jan.25,1896 p.4 neg	1
ff FXXX SX.X TXXx	p.41	1
ff F XXX SX.X T.Xx	p.50	
f F XXX SX GX	German Printing Speciment, p.31	1
F XXX SXX SXXx	v.2:19 neg	1
F XXX SXX JX	opp.p.76 neg	1
XX-		1
XXX-A		1
XX-A		20
C-EXXX	2	1
C-FXXX	v.2	1
C-BXX:16 neg 1		1

Table 4

Frequency of Occurrence of Location of Holding Field in
Bancroft Library Picture Collection Sample

Location of Holding	Frequency of Occurrence
picture drawer	175
picture	16
photo cabinet	6
Sierra Club Album	4
portrait drawer	3
Locked map case	1
picture cabinet	1

Collection Number

Collection Number is the number associated with the collection. The various forms of the contents of this field are illustrated in Table 5.

Gift

The Gift field contains information about the donor and/or the date of the item represented by a record in the present catalog. It is located in the lower left hand corner of records to which it applies.

Note

The Note field is the name given to miscellaneous information on a record that is not part of any other field. Miscellaneous notes are most often found in the lower left hand corner of the record.

See Reference

See Reference is the name given to the field containing a reference to another subject heading in the present catalog. This field always occurs in conjunction with the Subject Heading field.

Table 5
Various Forms of Collection Numbers

XX.X.XA ⁶
XX.X.XXA
XX.X.XXX
XX.X.XXXXA
A-X
A-AXXX
A-AXXXaa

⁶"A" represents a letter of the alphabet (upper case), "X" represents a number (e.g., 0, 1, 2, 3, 4, 5, 6, 7, 8, 9), and "a" represents a lower case letter of the alphabet.

PART II

File Conversion

The present catalog of the Bancroft Library Picture Collection must be converted to machine readable form in order to assemble a shelf list by computer. There are a number of possible methods by which this conversion could take place. One alternative is to type the contents of each card onto 8½ x 11" paper in a prescribed format using a standard typewriter. The sheets of paper could then be processed by an optical character reader and the information converted to machine readable form. The resulting records could then be edited by a computer and finally a book catalog of the collection could be produced.

A second alternative involves keypunching the records directly onto IBM cards and then machine processing the cards to produce a book catalog. The remaining part of this paper discusses a method employing the latter alternative.⁷

The conversion procedure has two parts: (1) Preparing the file for keypunching, and (2) Keypunching the information onto IBM cards according to a prescribed format.

File Preparation

As a prerequisite to computer processing of the cards, it is necessary to uniquely identify each card in each drawer of the Picture Catalog. It is also necessary to code each field on each card. The first step in accomplishing these two tasks is to photocopy every card in the present file. In this way,

⁷ Before selecting any of the possible alternative methods for conversion a cost study should be undertaken of the alternatives.

the records can be identified and coded without destroying or altering the original file.

The unique identification of each record entails numbering each card in the present file that contains a subject heading. Employing a numbering machine, or similar device, each card in the file that has a subject heading on it would be numbered consecutively--beginning with the first card in the first drawer and continuing through the last card in the last drawer. Only cards with subject headings would be numbered, and the number would be stamped in the upper right hand corner of the card.

The next step in file preparation is to code each field on each record. The suggested tagging codes are indicated in Table 6. The process of field coding involves writing the field code to the left of the field to which it applies. This process is best performed by a librarian or library assistant who is familiar with the present catalog and the field definitions.

An illustration of the numbering process and the coding process appears in Figure 6.

Record Layout

Once the file has been prepared according to the outlined procedure, then the actual conversion of the file to machine readable form can take place. The conversion will be accomplished by keypunching the contents of the fields of each record onto IBM cards.

Table 6
Tagging Codes

Code	Field
S	Subject Heading
D1	Description (1)
D2	Description (2)
N	Location Number
L	Location of Holding
T	Type of Holding
C	Collection
A	Collection Number
G	Gift
B	Note
R	See Reference

Figure 6 Part 1
Field Tagging and Numbering of
Cards in File Preparation

<p>1</p> <p>S Adobe House 205 Blithedale , Mill Valley C1 three exterior photos L picture drawer W 12991-12993</p> <p>G gift of Mrs. Lucretia H. Little 5.25.62 D enclosed in envelope are notes suggesting its origin.</p> <p>○</p>
<p>2</p> <p>S Carpenter, William Diarist D1 his sketches of places in Nicaragua and Mexico T photonegatives L picture drawer</p> <p>D2 Monkey Point N 17110(141)neg D2 Lake Nicaragua and Mts. N 17110(142)neg D2 San Juan del Sud harbor N 17110(143)neg D2 Cape Horn N 17110(144)neg D2 Acapulco Harbor N 17110(145)neg</p> <p>A all from C-B812 qv</p> <p>○</p>
<p>3</p> <p>S Carpinteria, California - Views D1 Giant Grape Vine trunk T photo N 11368 L picture drawer</p> <p>○</p>

Figure 6 Part 2
Field Tagging and Numbering of
Cards in File Preparation

S	Crater Lake, Oregon	41
R	see Gaytes Collection	
○		
S	Edouart, A. Artist	5
"	Headquarters of the Mendocino Indian Reservation	
T	Original	
C	Honeyman ^F 63.2.228C	
	Picture drawer	
	16903(402)4x5 color	
○		

The fields of each card will be keypunched in a prescribed sequence generally following the order in Table 7. That is, the first field punched on the card will be the Subject Heading, followed by See Reference, Description (1) Description (2), etc. If a certain field is not present on the record in hand the next applicable field should be punched. There will be no need to represent non-applicable fields. Each field has a varying number of characters in it. In order to signal the end of the field, a special delimiting character will be keypunched as the last character of the field.⁸

In addition to the field delimiter, each field will also be preceded by a tagging code. The tagging code is the same code assigned to a field in the file preparation. Codes will also be adopted to represent frequently occurring elements in certain fields. These codes can be found in Table 8.

A simple example of the keypunching format is given in Figure 7. The first six characters in Figure 7 comprise the primary number assigned to the subject heading "Carpenteria, California - Views." In columns 7 and 8 of the first card is the sequence number of the card. For the first card of Figure 7, the sequence number is 01 and for the second card the sequence number is 02. The first card of a new Subject Heading would have sequence number 01 with a different primary number.

The first field in Figure 7 is the Subject Heading, "Carpenteria, California - Views." It is preceded by its tagging code, S, in column 9. The field delimiter, #, directly follows the Subject Heading. There is no See Reference for this record so the next field punched is Description (1). The contents of Description (1) follow its tagging code, "D1." Since there is not enough room on the first IBM card to accommodate the field, it is continued onto a second card following the record number and the sequence number. It is followed by the field delimiter.

⁸The special character "#" is used in this paper as the delimiter. Any other suitable character could be substituted.

Table 7

Field Sequence for Key punching

Subject Heading

See Reference

Description (1)

Description (2)

Collection

Collection Number

Gift

Note

Type of Holding

Location of Holding

Location Number

Table 8
Field Codes

Code	Field	Explanation
P	Type of Holding	"Photo"
PS	Type of Holding	"Photos"
NG	Type of Holding	"Negative"
PN	Type of Holding	"Photonegative"
PNS	Type of Holding	"Photonegatives"
PD	Location	"Picture" or "Picture Drawer"
GO	Gift	"Gift of"
CL	Collection	"Collection"

Description (2), Collection, Collection Number, Gift, and Note are not applicable to this record, so the next field punched is the Type of Holding. The tagging code for this field, "T", is followed by a field code, "P", representing the contents "Photo." After the field delimiter, the next field punched is the Location of Holding. The tagging code for this field, "L", is followed by the field code "PD," representing picture drawer. The last field to be punched is the Location Number, 11368, preceded by its tagging code "N".

The record illustrated in Figure 7 is a simple example of the type of record found in the present catalog of the Bancroft Library Picture Collection. Many records in the present file have multiple occurrences of certain fields and certain sequences of fields. It becomes difficult to extrapolate from the simple example the keypunching format for more complex records.

Table 9 contains examples of possible sequences for the keypunching formats of complex records - i.e. records with more than one set of contents for a particular field. Subject Heading, See Reference, Collection, and Gifts were not found to repeat on the records sampled, so sample sequences for repetition of these fields have not been included.

The logic behind the field sequence examples in Table 9 is that the field sequence suggested in Table 7 should be followed. Fields with multiple contents are punched consecutively within the original sequence if they apply to the record as a whole. If the multiple contents of a field only apply to a portion of the record, then each will be punched within a new sequence along with the other fields to which they apply.

Table 9 contains examples of sequences in which multiple records apply to the record as a whole and to only a portion of the record. The second sample sequence in Table 9, S-D1-D2-T-L-N-D2-N is representative of the latter. This sequence represents the keypunching format for a record containing multiple Description (2)'s and corresponding Location Numbers. In Table 7,

Table 9
Sample Key punching Sequence for
Complex Records

Example Number	Sequence
1	S-D1-T-L-N-D1-N
2	S-D1-D2-T-L-N-D2-N
3	S-D1-C-A-T-L-N-T-L-N
4	S-B-B-T-L-N
5	S-D1-D2-B-T-L-N-B-T-L-N
6	S-D1-D2-G-T-L-N-N
7	S-D1-D2-C-A-T-L-N-A-N

Location Number is defined as the end of the field sequence for keypunching. The second of the two Description (2)'s and Location Numbers should then follow the first Location Number. Subsequent Description (2)'s and corresponding Location Numbers will follow this pattern.

This keypunching sequence is illustrated in Figure 8. The first sequence in the example begins with the Subject Heading, "Carpenter, William Diarist" and ends with the Location Number, 17110(141) neg. The second Description (2), "Lake Nicaragua and Mts.", and the second Location Number, 17110(142) neg, follow the first sequence. The computer will process this by reading in and storing the first sequence and then replacing the contents of Description (2) and Location Number with those entered in the second sequence. The rest of the Description (2)'s and their corresponding Location Numbers follow the same pattern.

The fourth sample sequence in Table 9, S-B-B-T-L-N, represents the sequence for a field whose multiple contents apply to the record as a whole. In this case, the two sets of contents of the Note field (B) are punched consecutively within the original sequence. This keypunching format is illustrated in Figure 9. The notes "photo of printing . . ." and "negative available . . ." both apply to the entire record so they are punched consecutively instead of in two different sequences.

Non-multiple occurring fields found in the original sequence do not always apply to each set of contents of multiple occurring fields found in later sequences. In this case, the tagging code for non-applicable fields should be punched by itself within each new sequence. A tagging code between two field delimiters will signal to the computer that that particular field as found in the original sequence does not apply to the portion found in the new sequence.

