"Reports of My Death Are Greatly Exaggerated"

Findings from the TEI in Libraries Survey

Michelle Dalmau, Digital Collections Services, Indiana University Libraries
Kevin S. Hawkins, Michigan Publishing, University of Michigan
4 October 2013, TEI Conference
Twitter: @mdalmau & @KevinSHawkins

Select milestones in the library digitization of text

- 1992: Founding of IATH at Virginia
- 1997: Founding of GDZ at SUB Göttingen
- 1999: Text Creation Partnership formed
- 2004: Google announced mass-digitization partnership with leading research libraries
- 2005: Open Content Alliance created
- 2008: Formation of HathiTrust
- 2010: TEI-C launches AccessTEI
- 2011: Release of a complete rewriting of the Best Practices for TEI in Libraries

Scope in brief
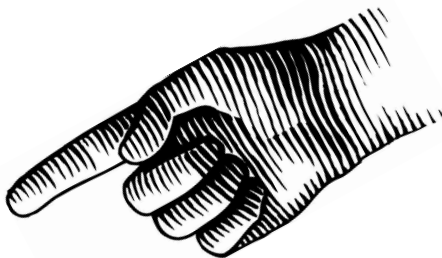
- Developed a SurveyMonkey survey with yes-no, multiple-choice, and free-response questions.

- Announced online on November 4, 2013 and closed on January 31, 2013.

- Survey participants had to answer "yes" to saying they work in a library.

- We encouraged responses from more than one person at the same institution.
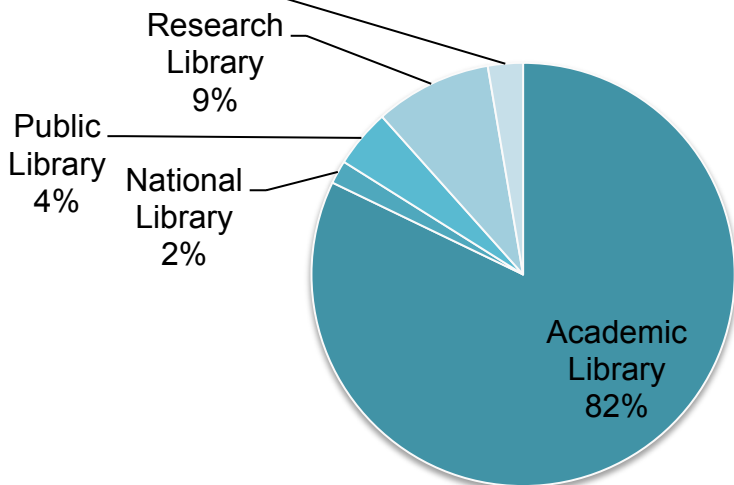
Who responded?

- 138 began the survey; 112 "completed" it
- We removed 26 responses from those who were supposed to have been disqualified for answering that they **did not work in a library**
- From the IP addresses of respondents, we can see that:
  - 55 are clearly affiliated with an institution; 41 of which are unique institutions
  - 57 are unidentifiable due to off-site internet connections (via ISPs)

Profile of Survey Respondents

Indicate the type of library for which you work.

Special Library 3%

Research Library 9%

Public Library 4%

National Library 2%

Academic Library 82%

Where is your institution located?

No Response 13%

Asia 2%

Europe 10%

North America 75%
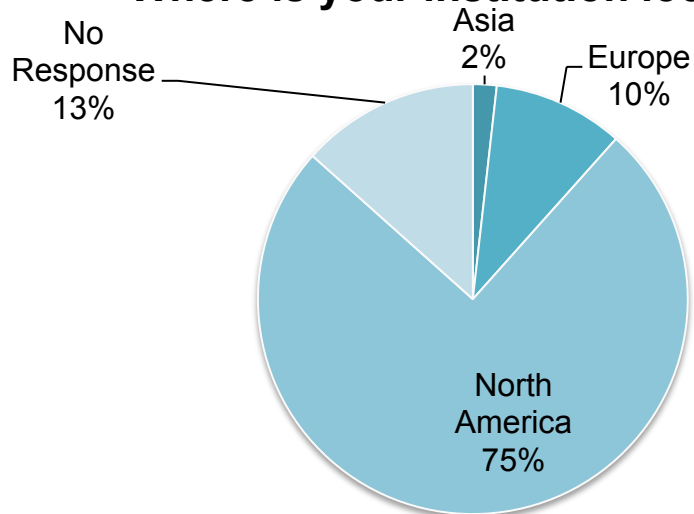
(n=112)

What is the size of your academic institution based on student enrollment *(or patrons served)*?

Up to 5,000 7%

5,000-10,000 14%

10,000-25,000 24%

25,000-40,000 20%

Over 40,000 13%
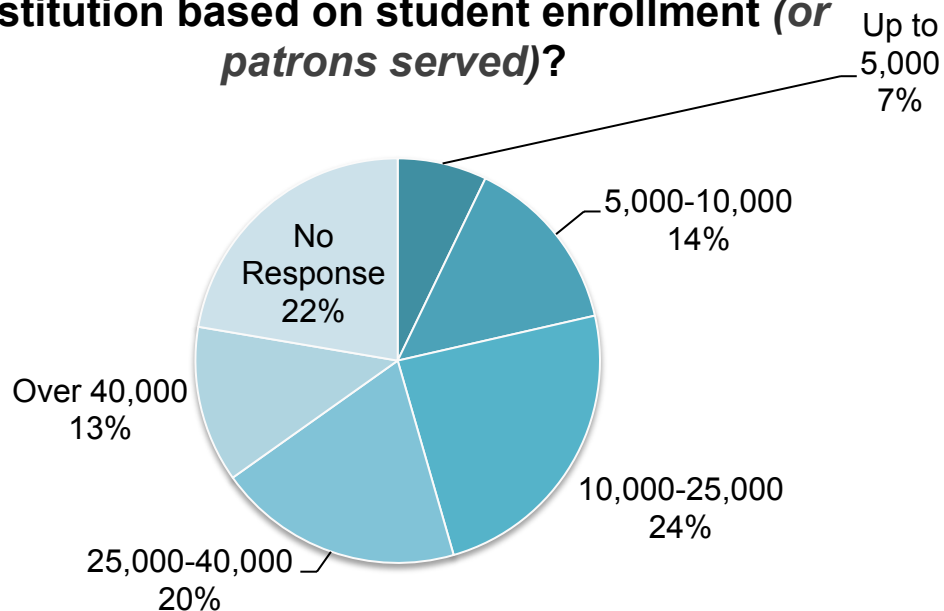
No Response 22%

What is the name of your unit or branch library?
(n=99 reported only one unit; n=9 reported more than one unit; n=4 no response)

- Administration 3%
- Archives 6%
- Cataloging 11%
- Collections 4%
- Digital Scholarship 5%
- General Library 14%
- Preservation 1%
- Public Services 2%
- Publishing 3%
- Special Collections 11%
- Subject Library 8%
- Technology 32%
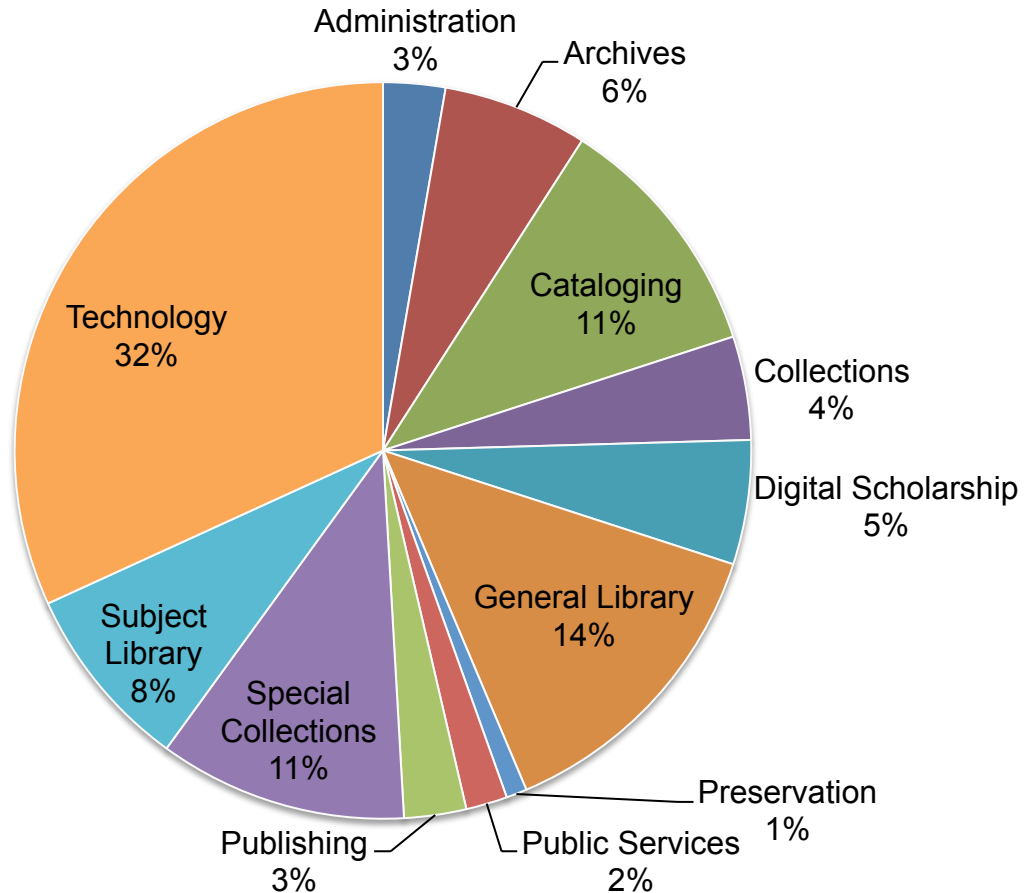
List the units or people (in terms of roles) in your library with which/ whom you partner (n=58)
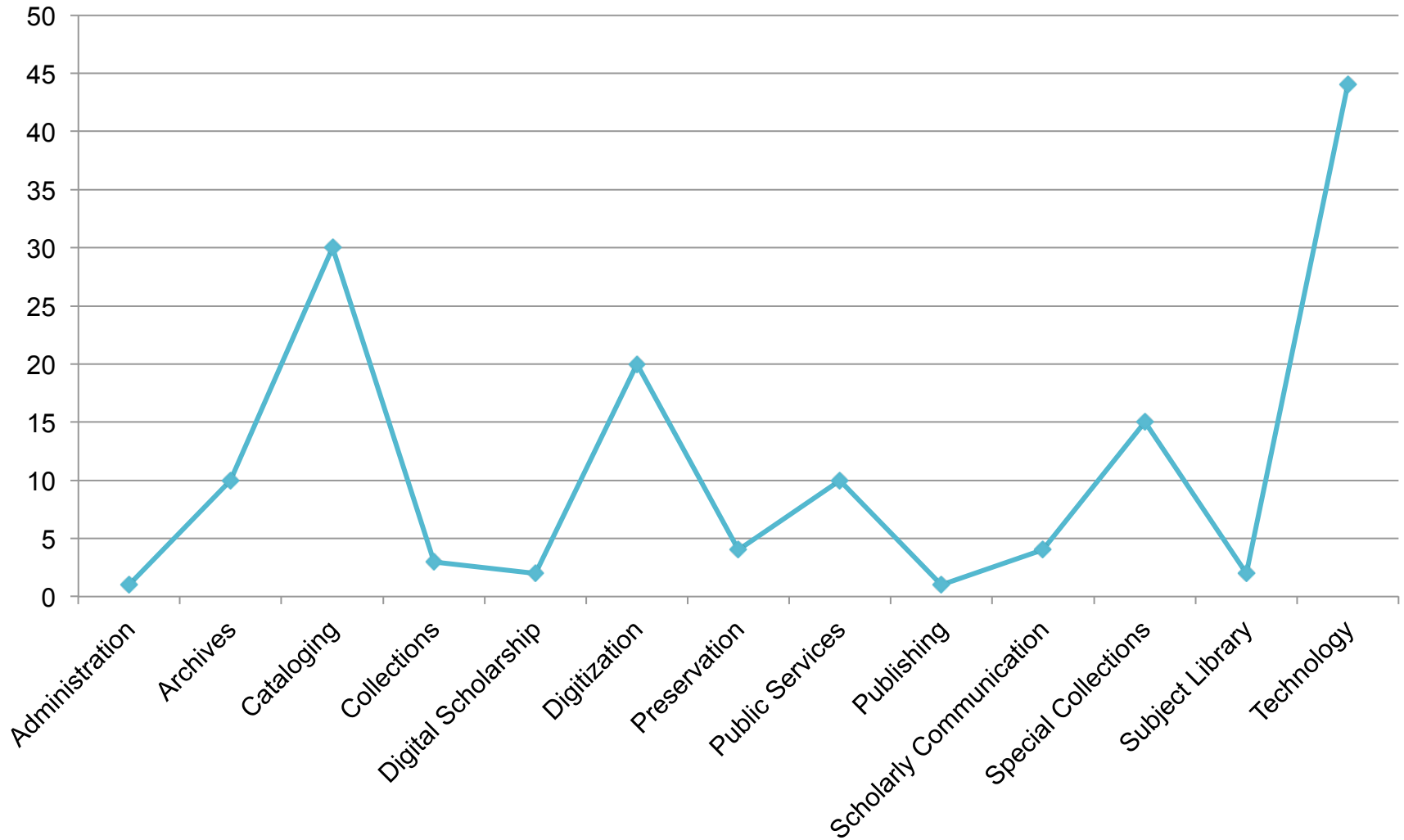
TEI-C Membership

Is your institution a member of the TEI Consortium? (n=112)

Legend: ■ Total Responses ■ Total Institutions ■ Total Unique Institutions ■ Total ISPs

	Total Responses	Total Institutions	Total Unique Institutions	Total ISPs
Yes	39	18	9	21
No	43	23	17	20
Unsure	7	5	6	2
No Response	23	9	9	14
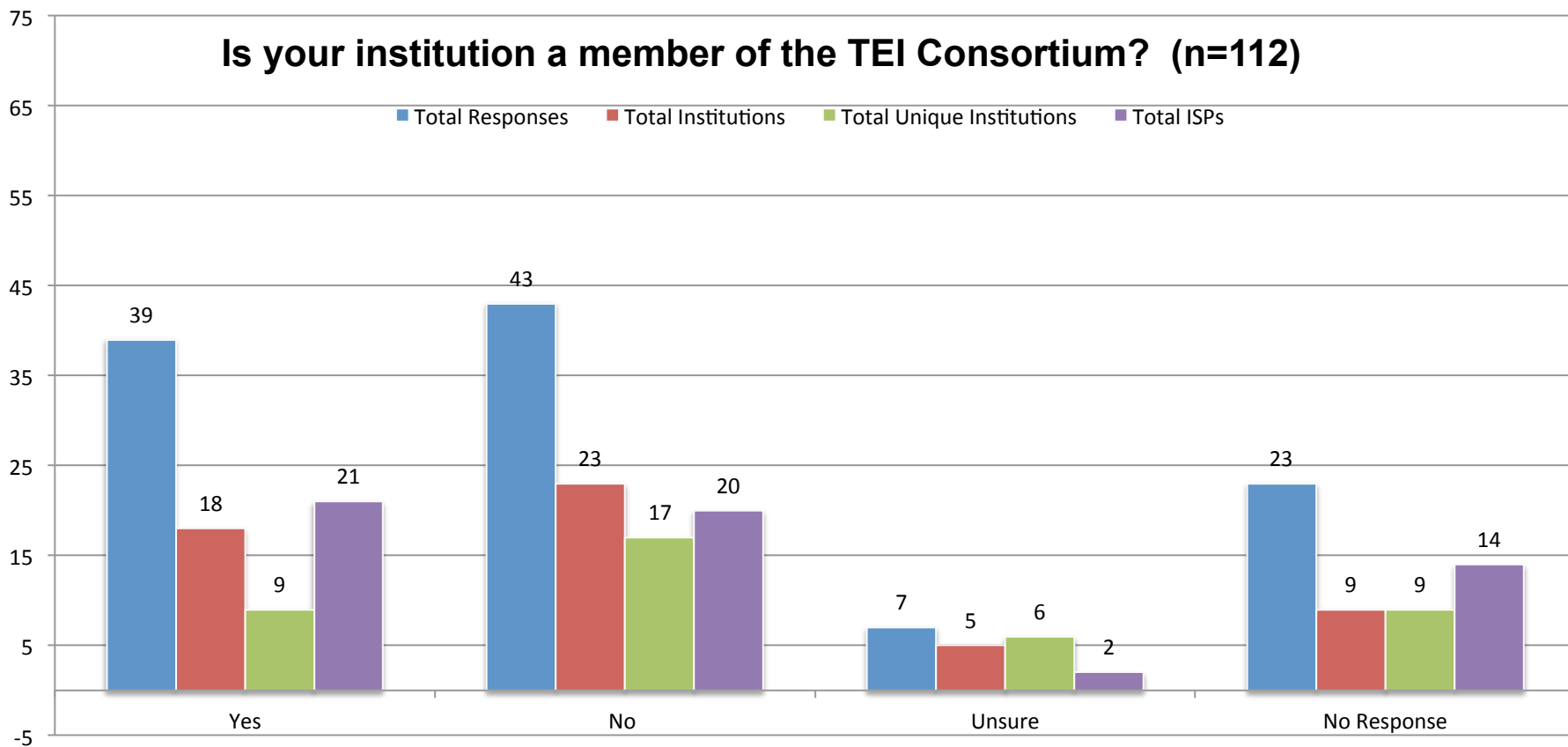
TEI-C Member Institutions: 2005-2013
(missing 2012 data)

■ Libraries ■ Non-Libraries ■ Combination ■ Unsure
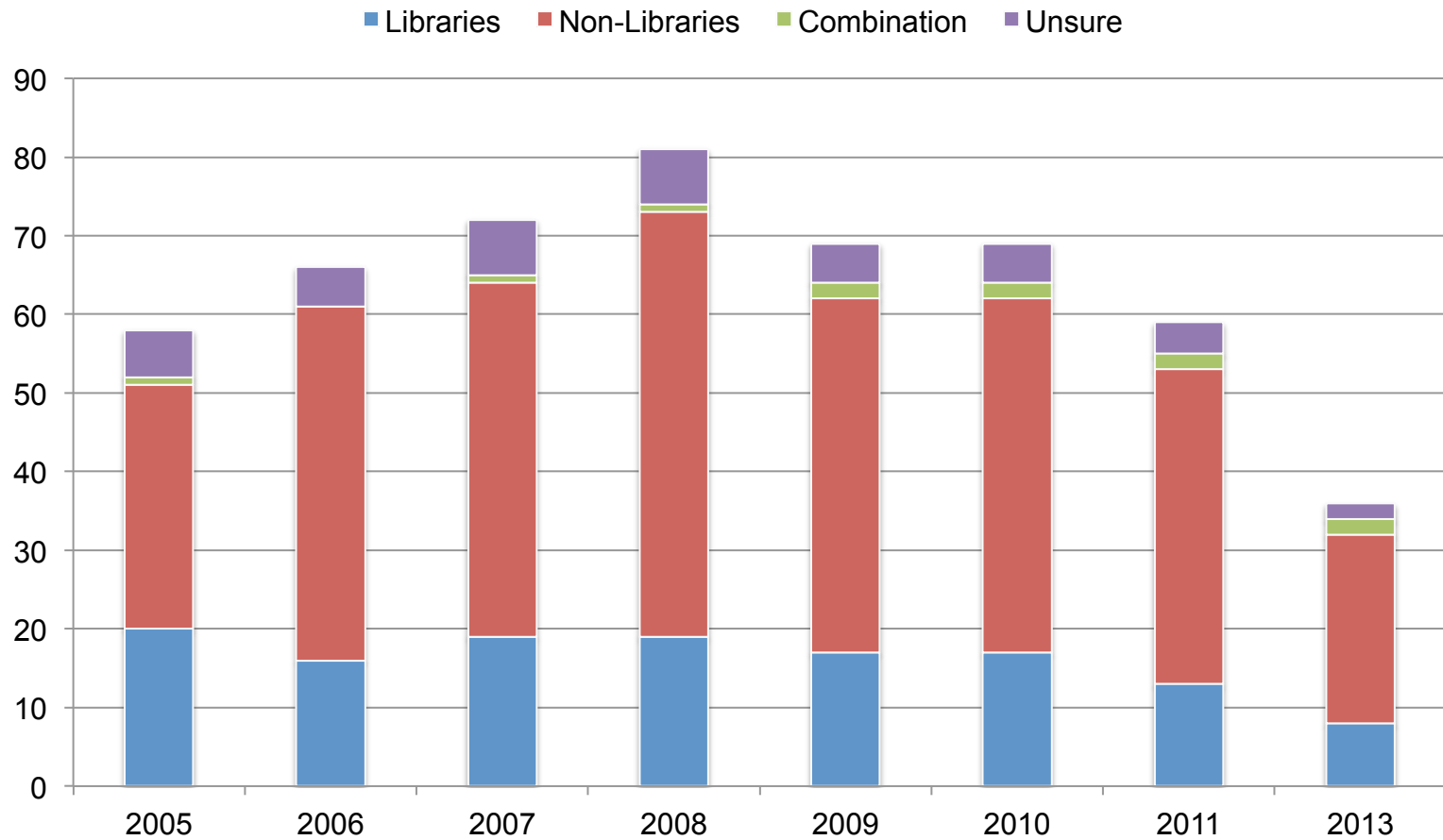
*Membership data (2005-2011) provided by Martin Mueller; Coded by Kevin Hawkins

Would your institution become or remain a TEI Consortium member if doing so would allow for: (n=112)

Legend: ■ Number of Responses ■ No response

Category	Number of Responses	No response
Digitization discounts	25	87
Direct input into standards development	20	92
Community support	31	81
Training opportunities	43	69
Other	18	94

Pie chart:
- 3% belong regardless of benefits
- 1% consulting opportunities
- 1% tool development
- 3% no incentives possible
- 92% no responses
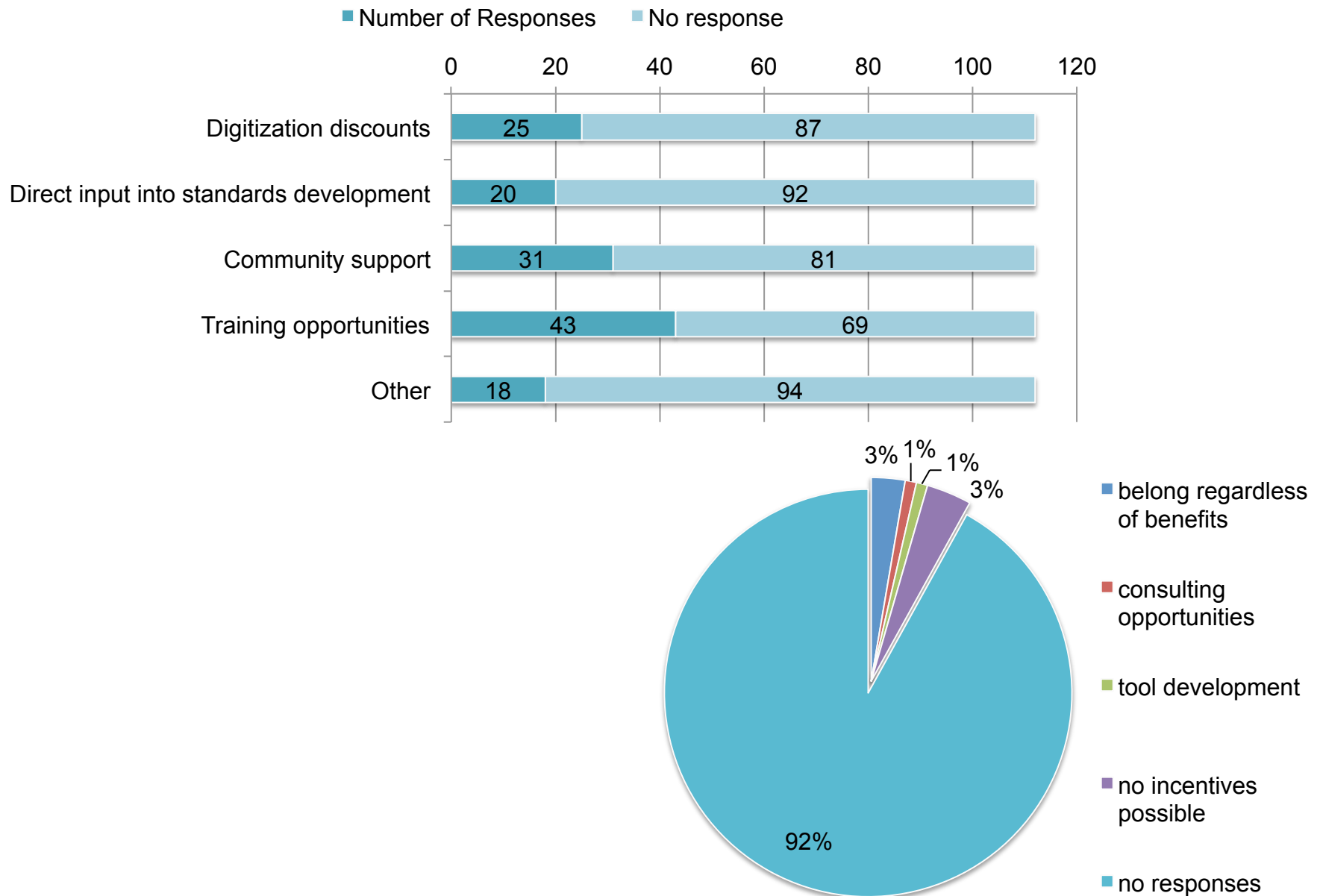
Text Encoding Practices & Partnerships in Libraries

How many text encoding projects have you or members of your unit participated in over the years? (n=112)

None
4%

1-5
40%

6-10
13%

11-15
7%

16-20
2%

More than 30
12%

No Response
22%
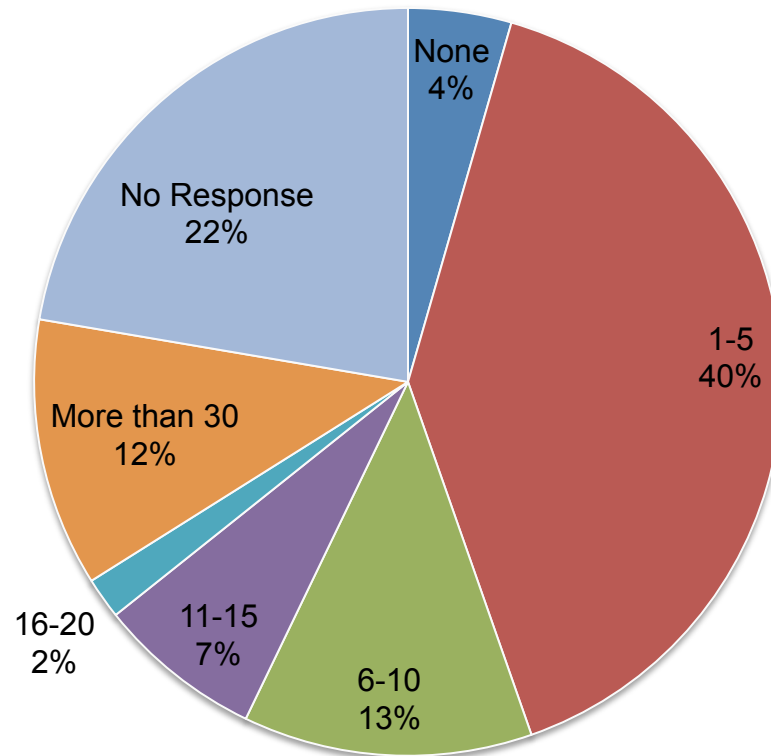
In what ways do you or members of your unit support text encoding projects (select all that apply)? (n=112)

■ Number of Responses ■ No response

Category	Number of Responses	No response
Consulting	44	68
Training	44	68
Project management	49	63
Establishing encoding workflows	48	64
Text encoding/markup	61	51
Web publishing	43	69
Other	7	105
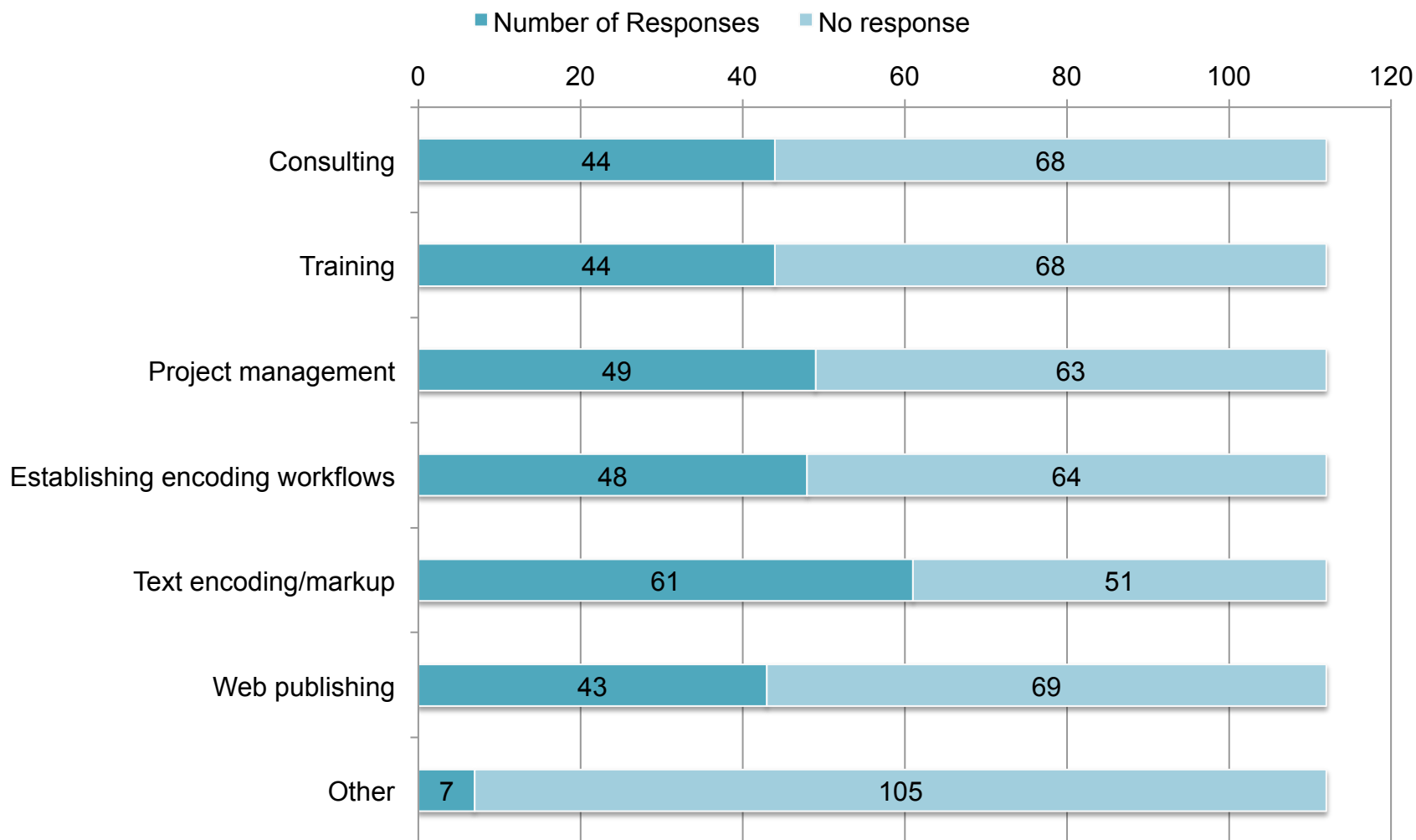
Indicate how frequently you partner with the following people when undertaking a text encoding project:

■ Almost Always/Often   ■ Sometimes   ■ Seldom/Never

	Almost Always/Often	Sometimes	Seldom/Never
IT Professionals	33	20	15
Museum Professionals	4	3	58
Archivists	22	16	28
Librarians & Library Staff	37	13	17
Graduate Students	19	22	27
Faculty	33	17	19
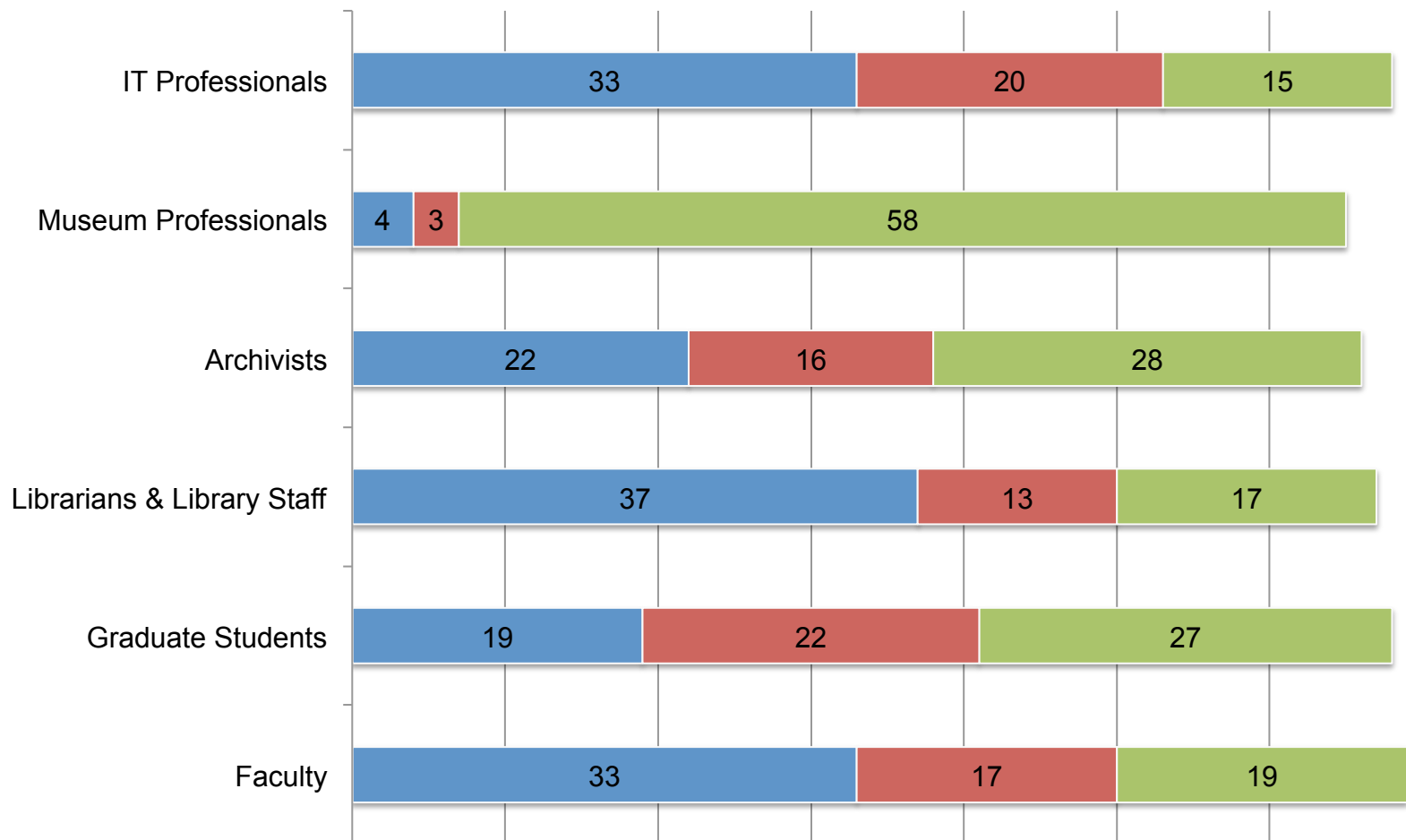
Rank the nature of your text encoding projects by "dragging and dropping" each option into place (1 is most common, 8 is least common)

- Rare Books & Manuscripts
- Archival Materials
- Faculty or Librarian Digital Research Projects
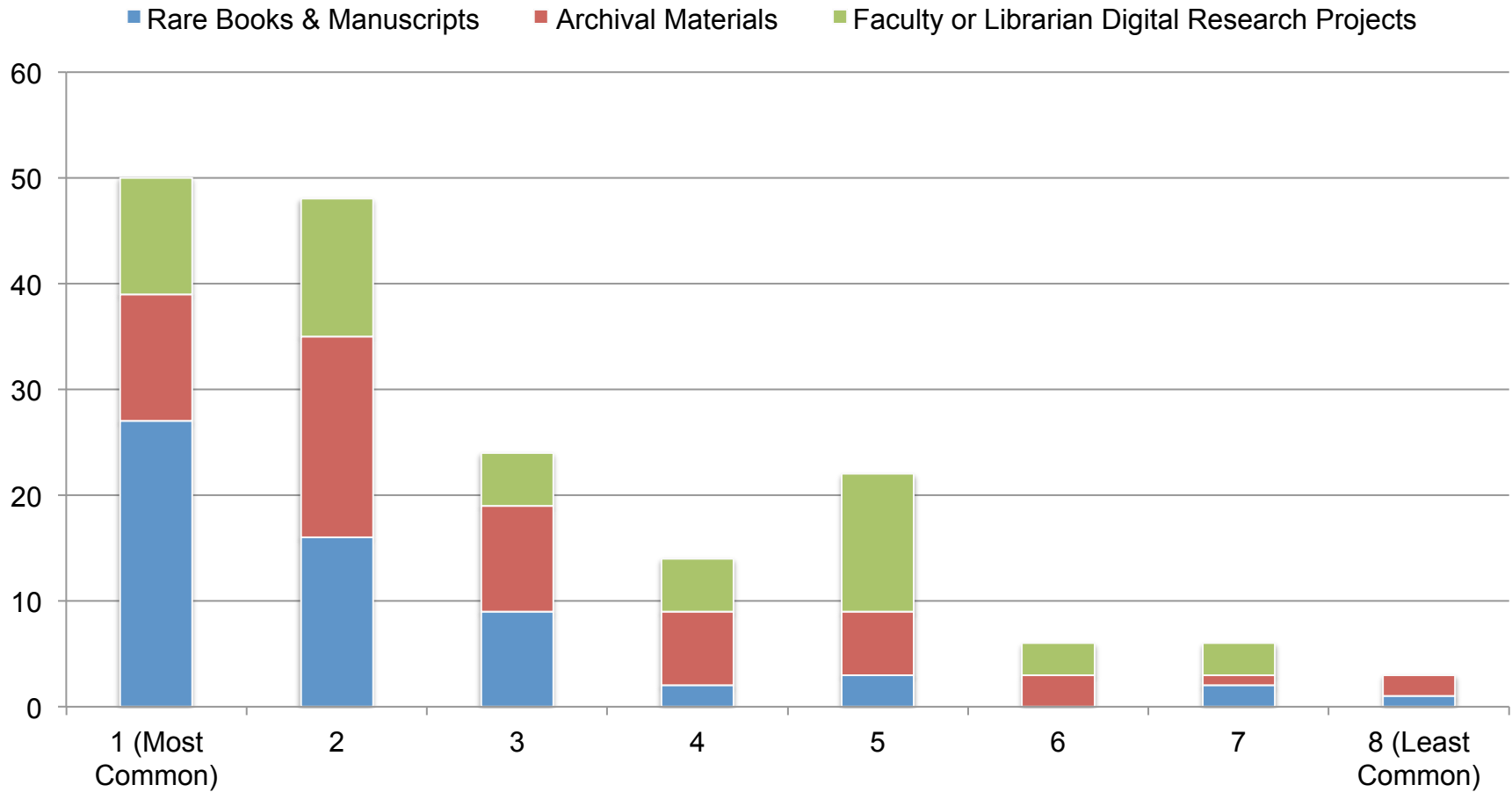
What type of materials do you digitize and encode? Indicate frequency.

Legend: ■ Almost Always/Often ■ Sometimes ■ Seldom/Never

Material	Almost Always/Often	Sometimes	Seldom/Never
Manuscripts / Handwritten Materials	35	16	15
Printed Books	35	14	17
Transcriptions of Audio or Video	6	7	43
Born-Digital Works	9	12	38
Catalogs of Manuscripts or Other Items	5	7	42
Newspapers	5	11	44
Serials	9	13	40

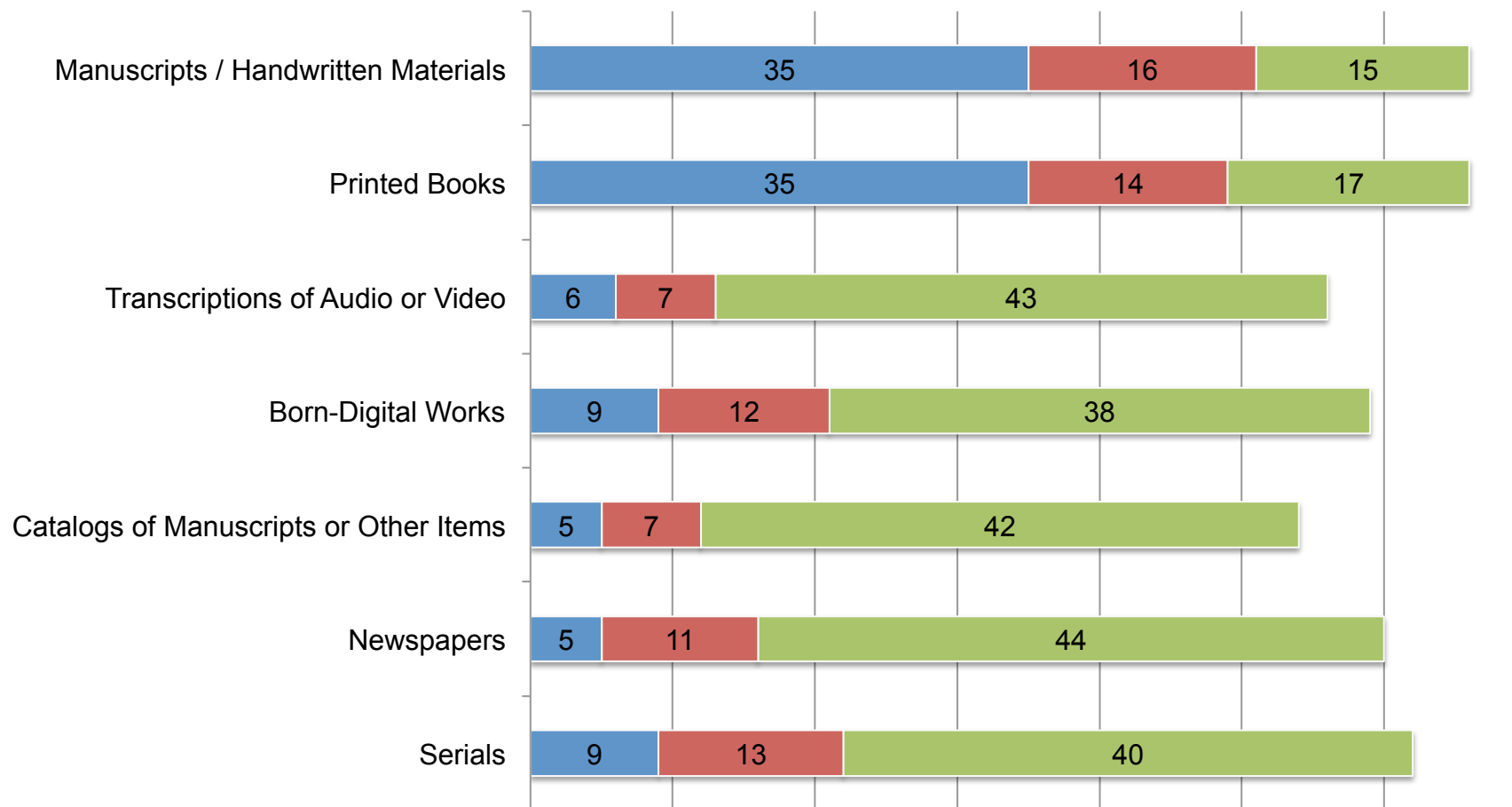# Indicate the frequency in which materials from the following time periods are encoded:

■ Almost Always/Often ■ Sometimes ■ Seldom/Never

Time period	Almost Always/Often	Sometimes	Seldom/Never
19th century and after	43	15	8
18th century	23	17	17
17th century	18	11	23
16th century	14	9	27
15th century and before	11	6	29

Describe the types and frequency of encoding projects you undertake based on the following levels of encoding:

■ Almost Always/Often ■ Sometimes ■ Seldom/Never

Level of encoding	Almost Always/Often	Sometimes	Seldom/Never
Basic reformattiing of text	27	7	19
Mid-level structural encoding	36	14	16
Richer encoding for content analysis	22	18	22
Scholarly encoding projects	25	15	22

Levels of Encoding (Q89-Q91) v. Number of Encoding Projects (Q24)

Legend: 0 — 1-5 — 6-10 — 11-15 — 16-20 — More than 30

Basic Reformatting of Text (Q89): Almost Always, Often, Sometimes, Seldom, Never
Mid-level Structural Encoding (Q90): Almost Always, Often, Sometimes, Seldom, Never
Richer Encoding for Content Analysis (Q91): Almost Always, Often, Sometimes, Seldom, Never
Scholarly Encoding Projects (Q92): Almost Always, Often, Sometimes, Seldom, Never

Levels of Encoding (Q89-Q91) v. Number of Encoding Projects (Q24)
(n=40)

Legend: 1-5 | 6-10 | More than 30

Basic Reformatting of Text (Q89)
Mid-level Structural Encoding (Q90)
Richer Encoding for Content Analysis (Q91)
Scholarly Encoding Projects (Q92)

Categories: Almost Always, Often, Sometimes, Seldom, Never

Level of Encoding v. Types of Materials

Legend:
- Manuscripts
- Printed Books
- Transcriptions of Audio/Video
- Born-digital Works
- Catalogs of Manuscripts
- Newspapers
- Serials

X-axis categories:
- Basic Reformatting of Text (Q89): Almost Always/Often, Sometimes, Seldom/Never
- Mid-level Structural Encoding (Q90): Almost Always/Often, Sometimes, Seldom/Never
- Richer Encoding for Content Analysis (Q91): Almost Always/Often, Sometimes, Seldom/Never
- Scholarly Encoding Projects (Q92): Almost Always/Often, Sometimes, Seldom/Never

Level of Encoding v. Types of Materials (top two items)

Manuscripts — **Printed Books**

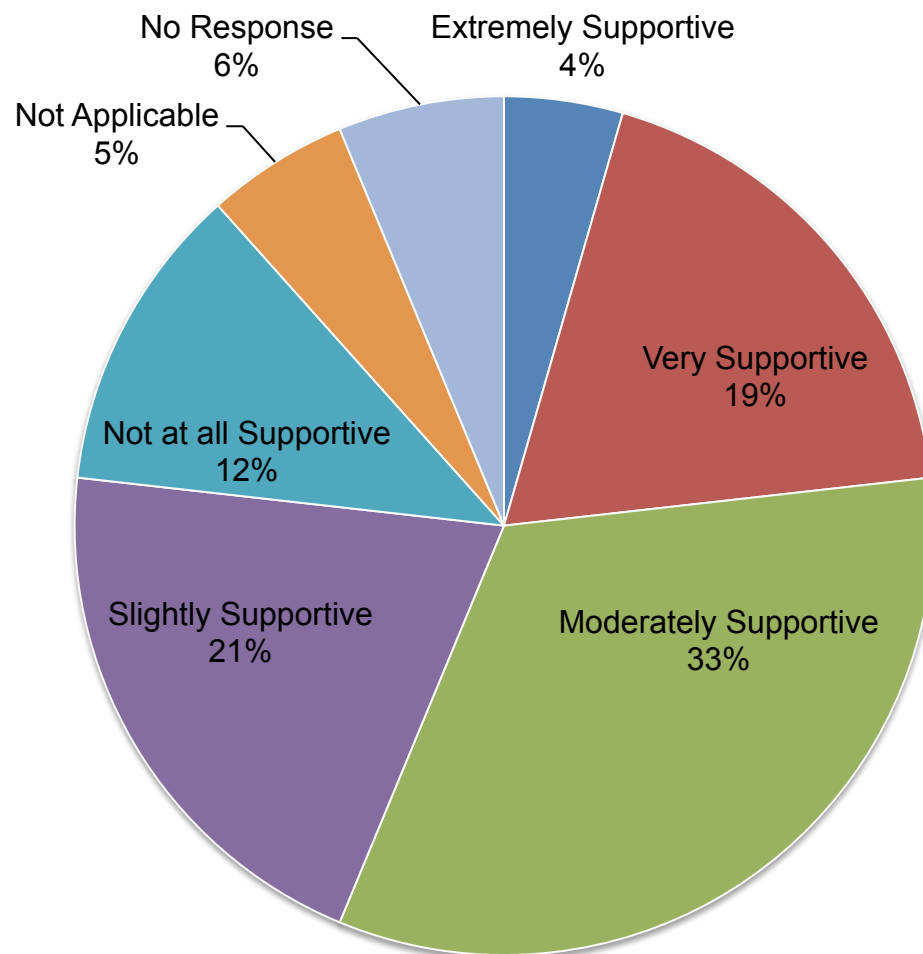|	Basic Reformatting of Text (Q89)			Mid-level Structural Encoding (Q90)			Richer Encoding for Content Analysis (Q91)			Scholarly Encoding Projects (Q92)		
Category	Almost Always/Often	Sometimes	Seldom/Never	Almost Always/Often	Sometimes	Seldom/Never	Almost Always/Often	Sometimes	Seldom/Never	Almost Always/Often	Sometimes	Seldom/Never
Manuscripts	14	10	8	18	10	6	15	9	7	14	8	8
Printed Books	17	11	4	25	5	3	13	10	9	15	9	9

Text Encoding Interests & Attitudes in Libraries

How would you rate the level of interest in text encoding by members of your library as whole? (n=112)

- No Response 5%
- Extremely Interested 1%
- Not Applicable 4%
- Very Interested 9%
- Not at all Interested 14%
- Moderately Interested 32%
- Slightly Interested 35%
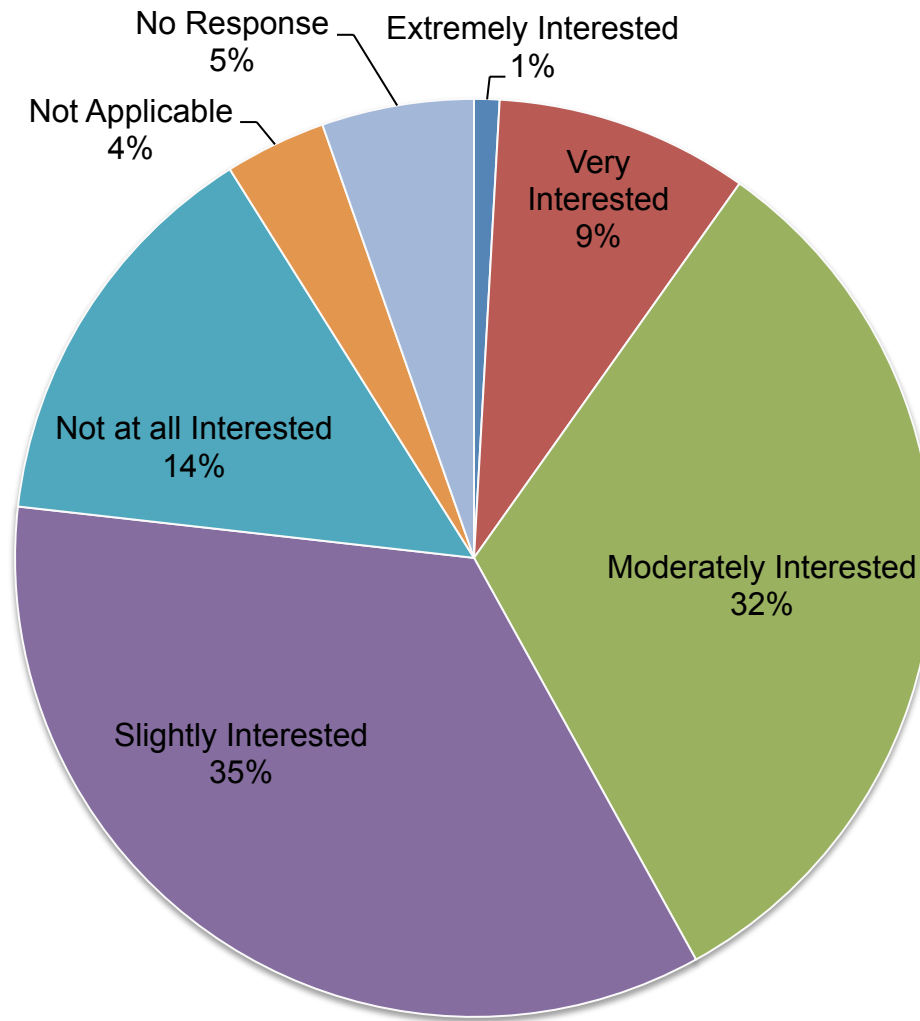
How would you rate your library's administrative support for text encoding projects today? (n=112)

No Response
6%

Extremely Supportive
4%

Not Applicable
5%

Very Supportive
19%

Not at all Supportive
12%

Moderately Supportive
33%

Slightly Supportive
21%

Q5 * Q6 Crosstabulation

Count

		Q6			
		Extremely Interested	Moderately Interested	N/A	Not at all Interested
Q5		6	0	0	0
	Extremely Supportive	0	1	3	0	1
	Moderately Supportive	0	0	17	1	3
	N/A	0	0	0	3	0
	Not at all Supportive	0	0	3	0	5
	Slightly Supportive	0	0	4	0	4
	Very Supportive	0	0	9	0	3
Total		6	1	36	4	16

Q5 * Q6 Crosstabulation

Count

		Q6		Total
		Slightly Interested	Very Interested	
Q5		1	0	7
	Extremely Supportive	0	0	5
	Moderately Supportive	13	3	37
	N/A	3	0	6
	Not at all Supportive	5	0	13
	Slightly Supportive	13	2	23
	Very Supportive	4	5	21
Total		39	10	112

Legend: ■ Extremely Supportive ■ Very Supportive ■ Moderately Supportive ■ Slightly Supportive ■ Not Supportive
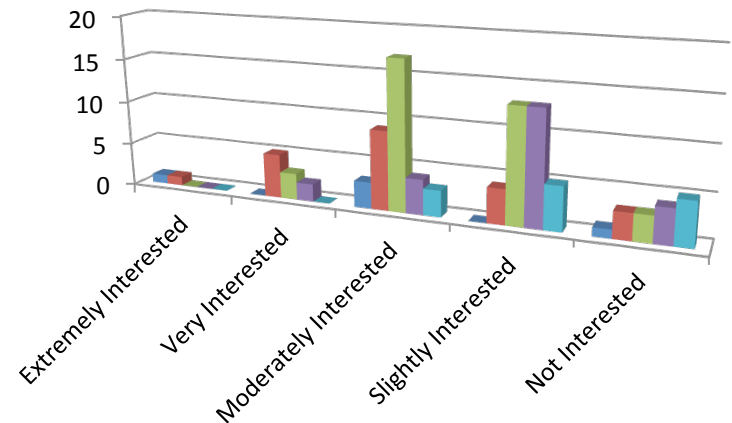
In a few sentences, could you describe how you see the state of and attitudes toward text encoding in your library today? (n=63)
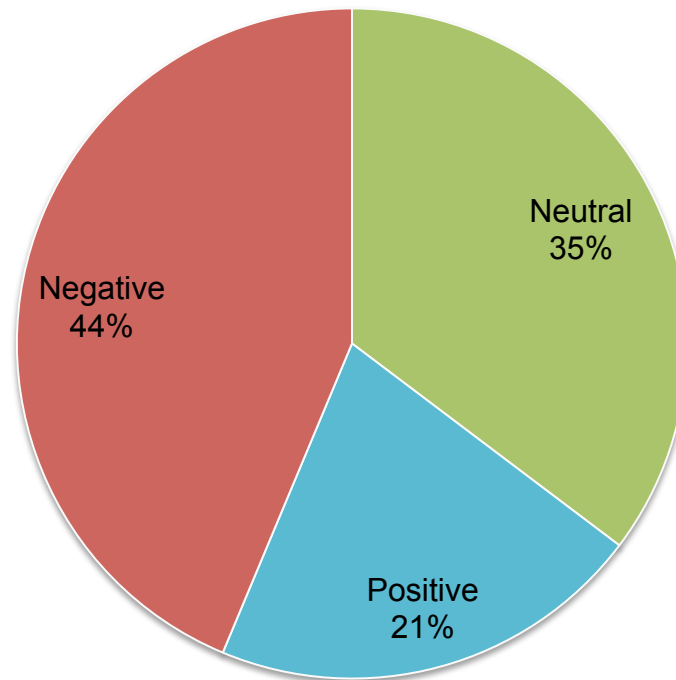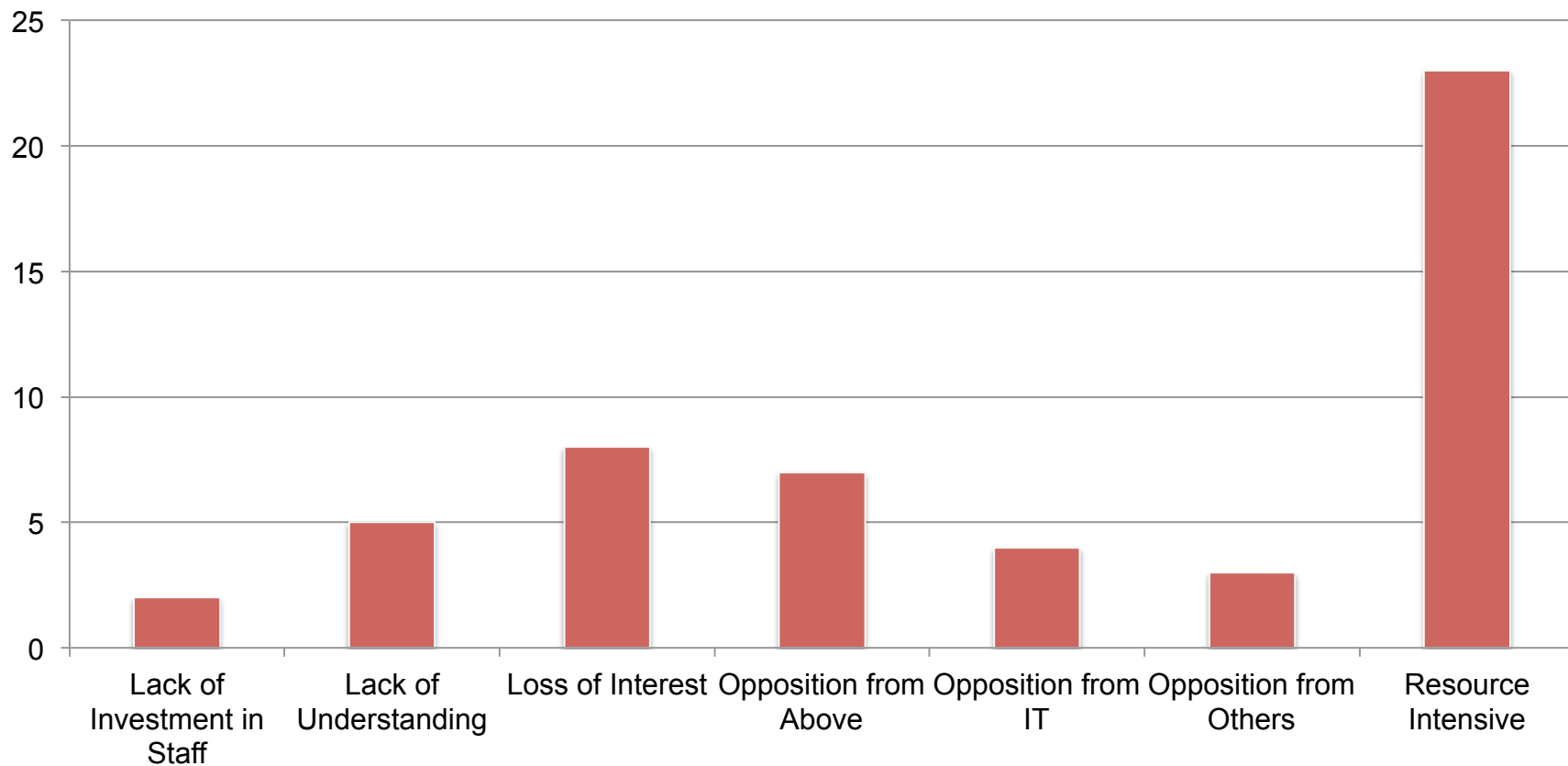
Neutral
35%

Negative
44%

Positive
21%

In a few sentences, could you describe how you see the state of and attitudes toward text encoding in your library today?
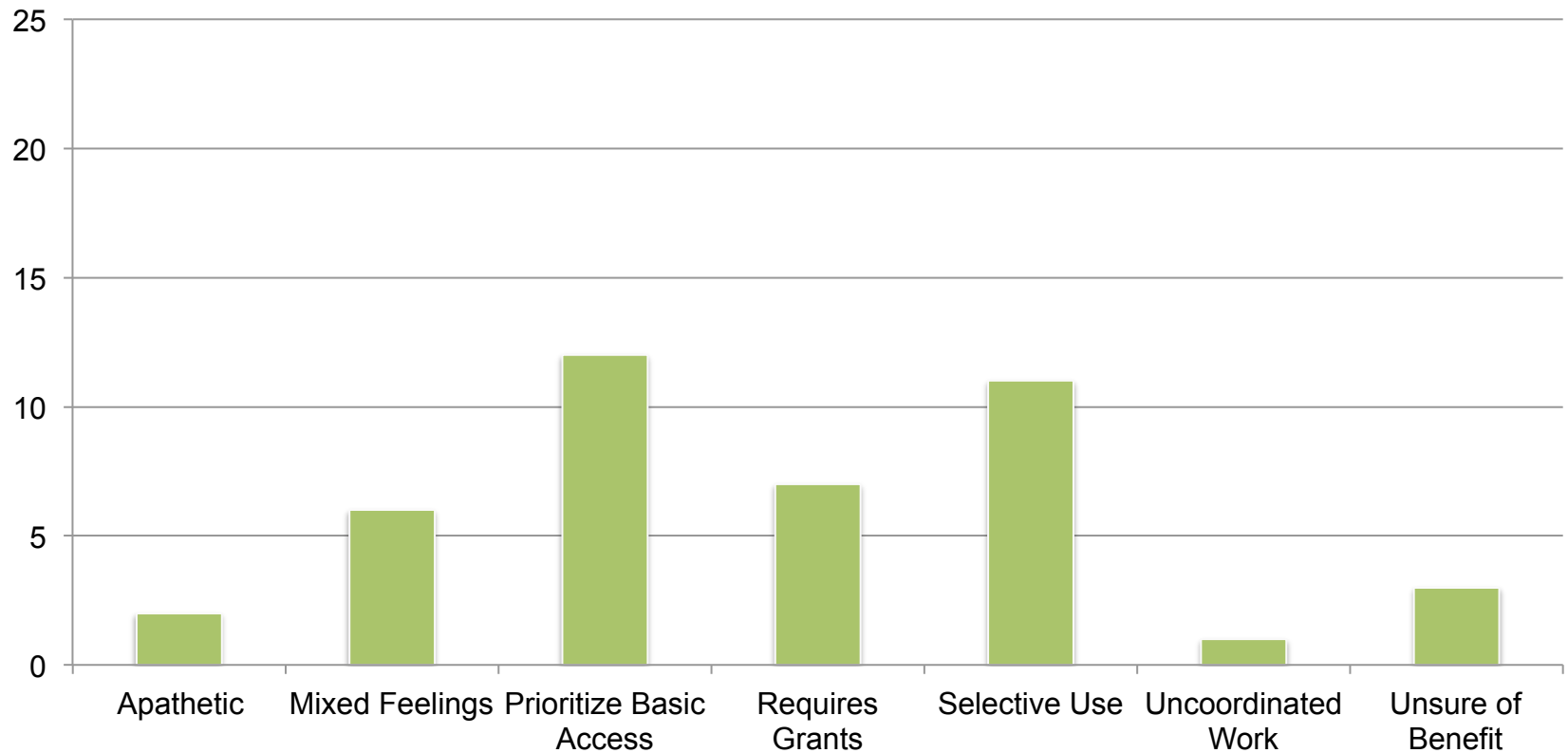## Coded Positive
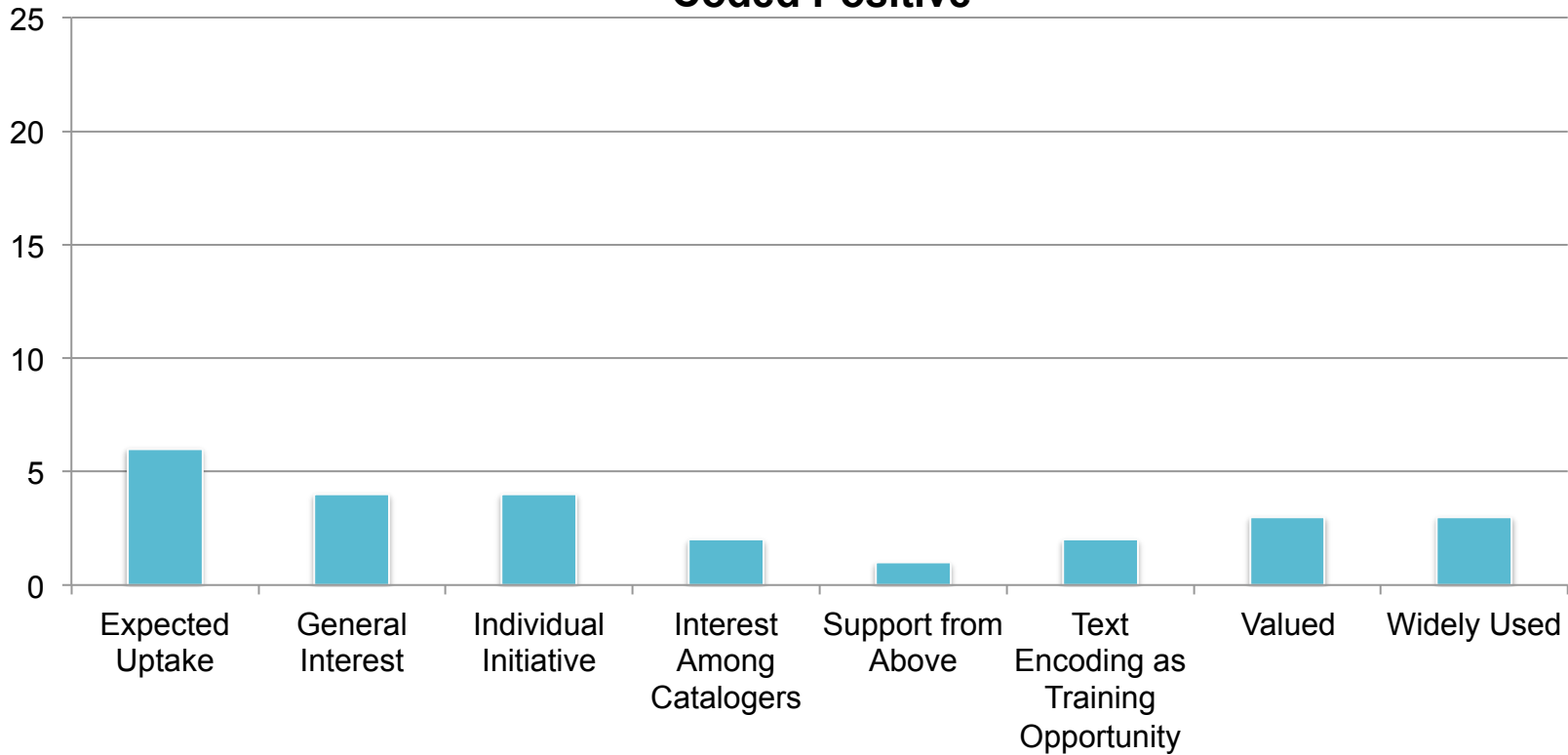
In a few sentences, could you describe how you see the state of and attitudes toward text encoding in your library today?
Coded Negative

In a few sentences, could you describe how you see the state of and attitudes toward text encoding in your library today?
Coded Neutral

What's Next?

You Ask.