

Research data quality assurance at repositories indexed in re3data

METHOD

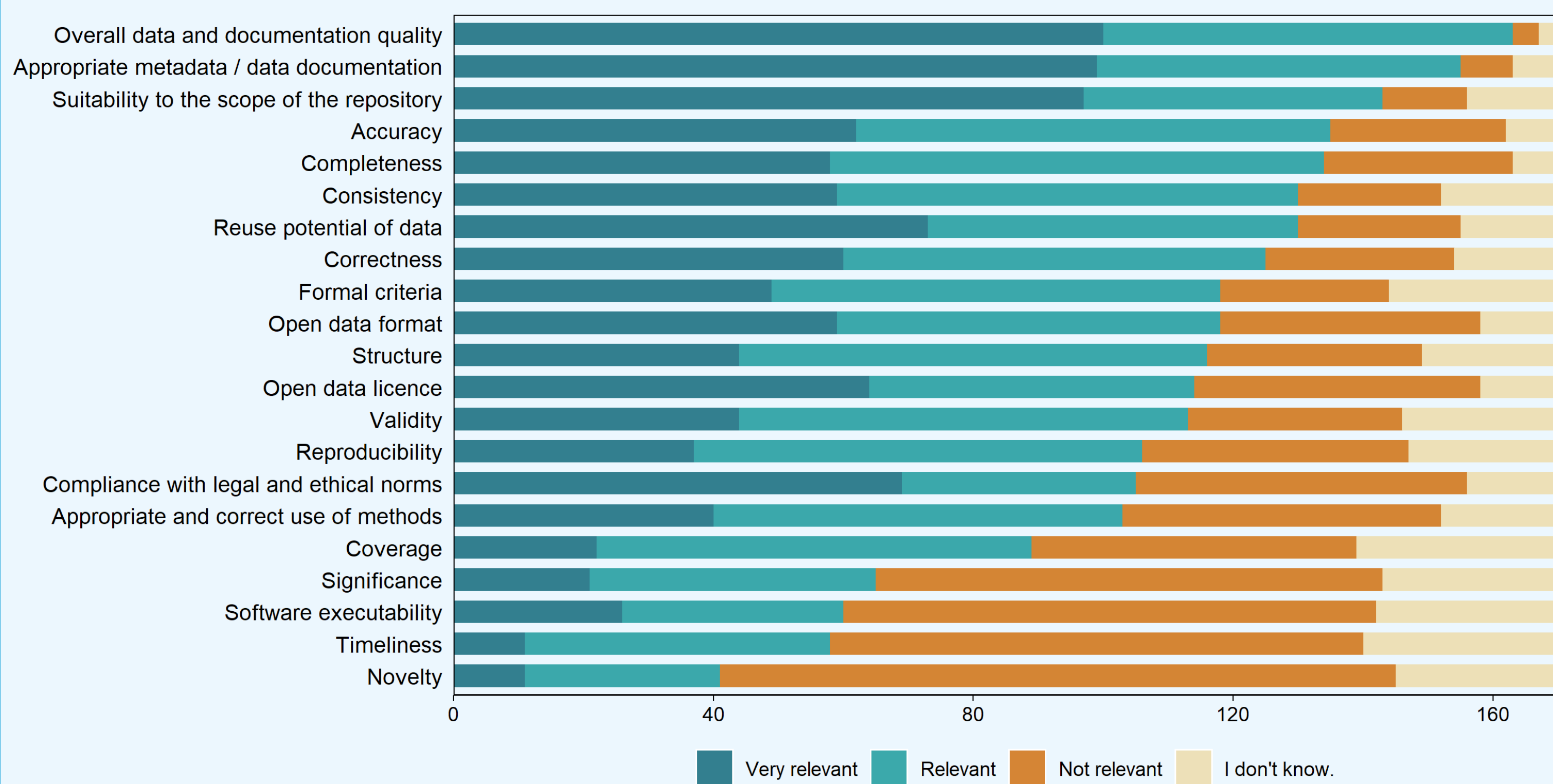
Quality assurance is a central challenge when sharing research data, as it ensures that data are valid, reliable, and usable. The landscape of repositories and their essential contribution to research data sharing is well studied. In contrast, we know much less about repositories' roles in research data quality assurance, and their contributions remain largely invisible.

To address this issue, we conducted a survey among staff responsible for the data curation at repositories listed in re3data, an international registry of research data repositories. Of the 1897 repositories that were contacted, 332 completed the questionnaire.

1. QUALITY INDICATORS

How relevant are the following quality criteria for data review at your repository? (N=171)

Figure 1

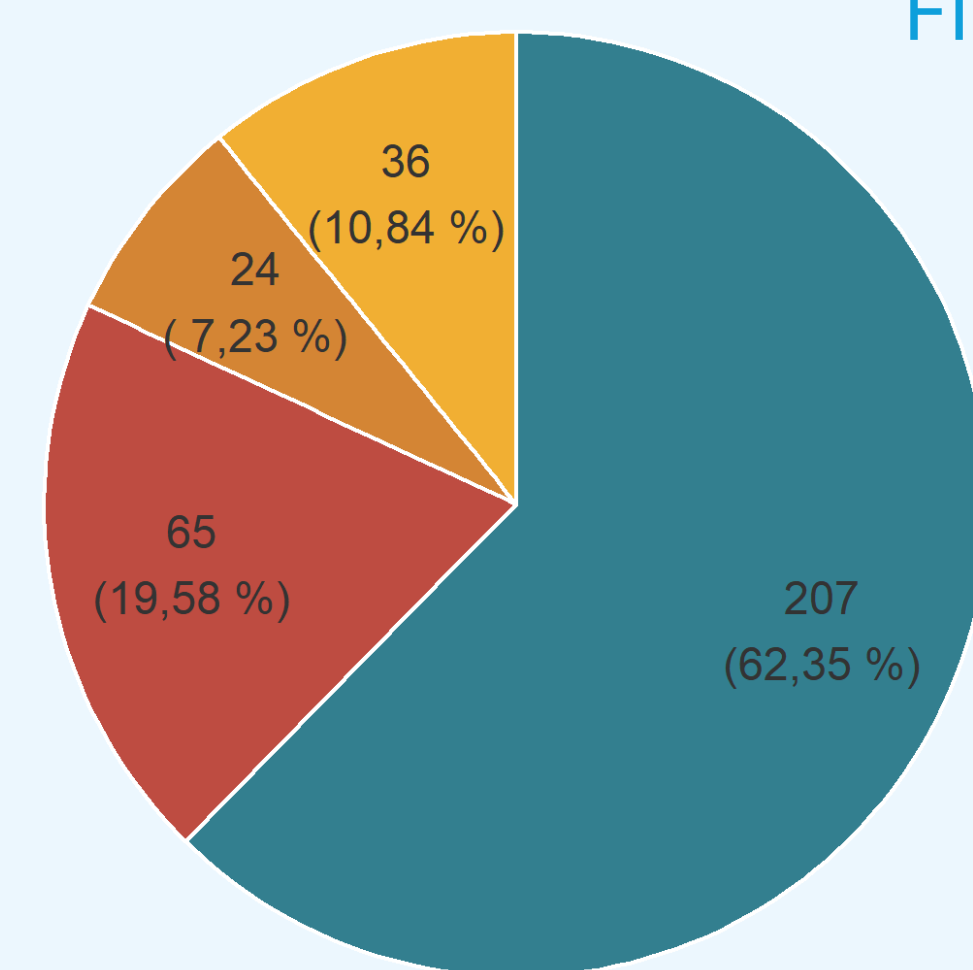


2. FORMAL ASSESSMENT

Are formal criteria applied to data before publication? (N=332)

Figure 2

"Formal criteria" refers to technical, administrative, and access-related aspects of data to be published.



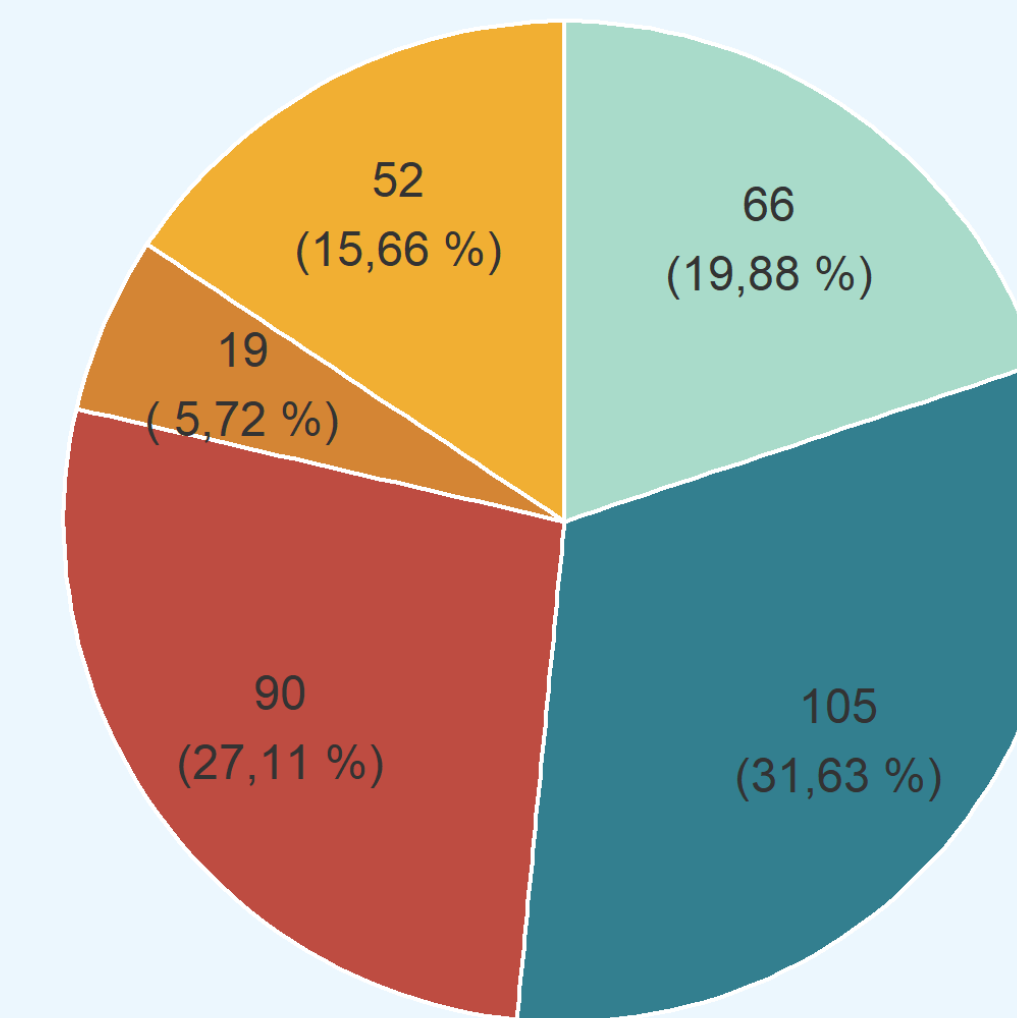
Legend for Figure 2:
 - Yes, formal criteria are applied. (Dark Blue)
 - No, formal criteria are not applied. (Red)
 - I don't know. (Orange)
 - Formal criteria are not applicable to the repository. (Yellow)

3. DATA REVIEW

Are data reviewed beyond the application of formal criteria? (N=332)

Figure 4

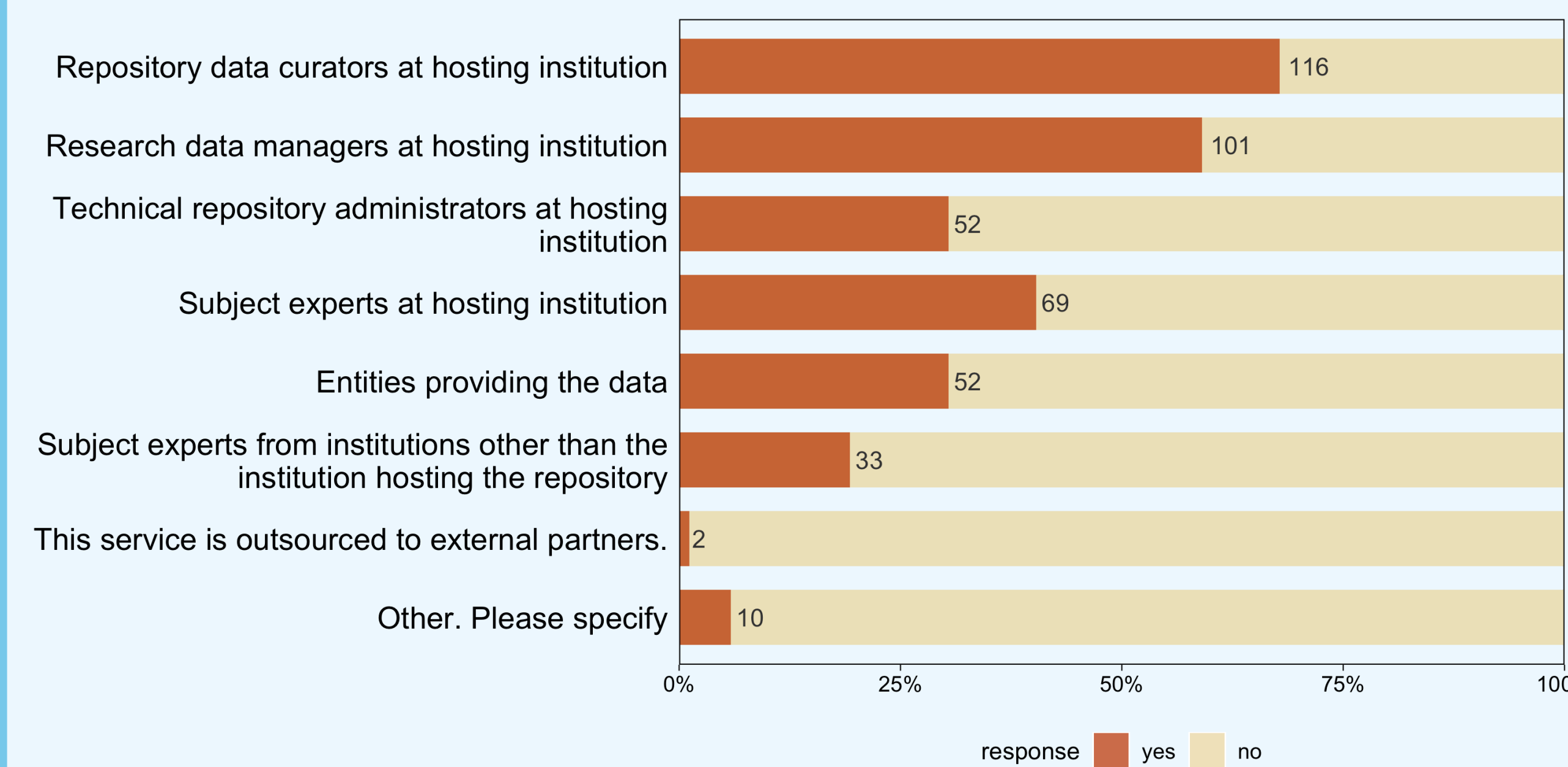
"Data review" refers to the process by which experts, either from the hosting institution or other institutions, evaluate the "scientific" quality of datasets, such as methods or reusability.



Legend for Figure 4:
 - Yes, certain datasets. (Light Green)
 - Yes, all datasets. (Dark Blue)
 - No (Red)
 - I don't know. (Orange)
 - Data review is not applicable to the repository. (Yellow)

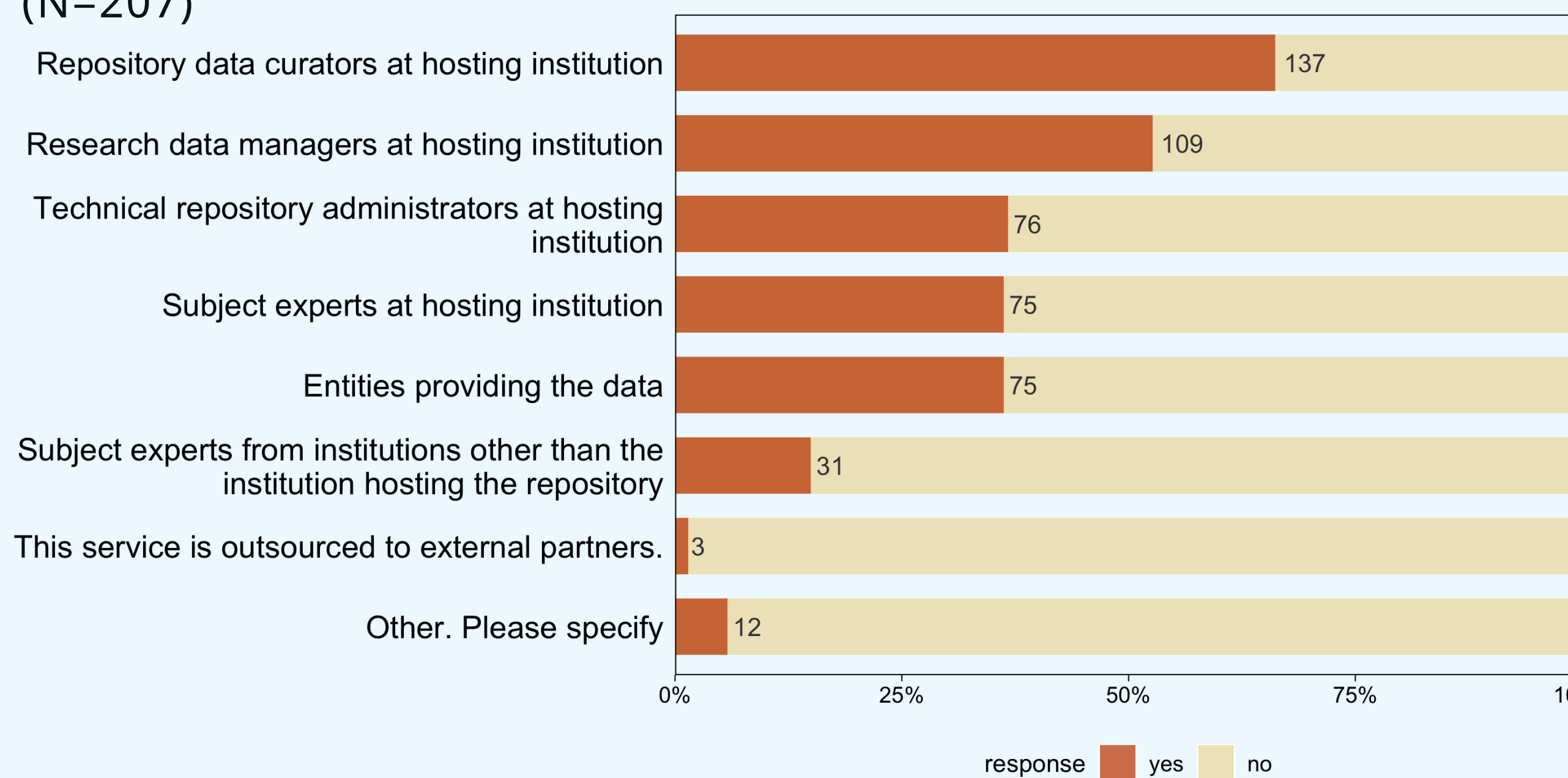
Who is responsible for data review? (N=171)

Figure 5



Who is responsible for data curation at your repository? (N=207)

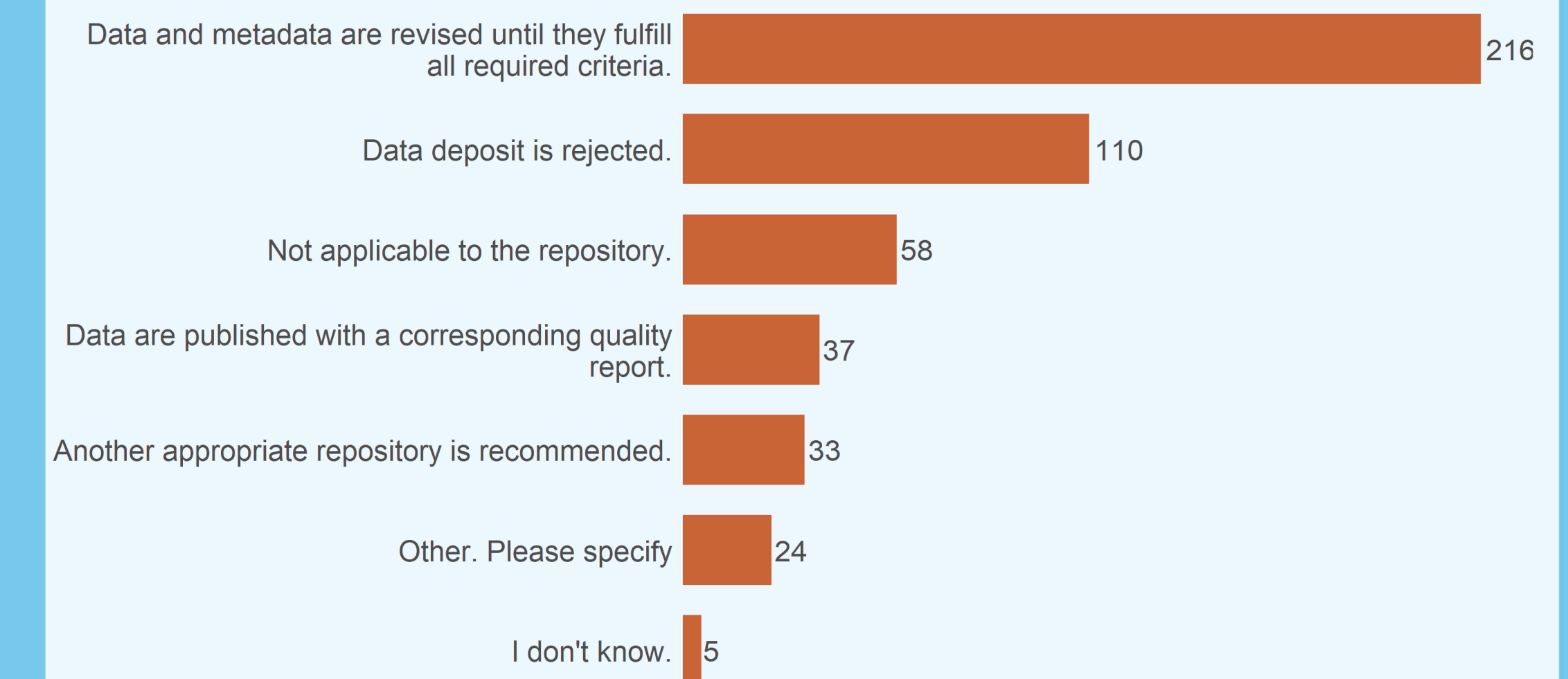
Figure 3



4. DATA REJECTION

What are the consequences of submitting data of insufficient quality? (N=332)

Figure 6



What (estimated) ratio of datasets were rejected by your repository in the last two years? (N=86)

Figure 7

