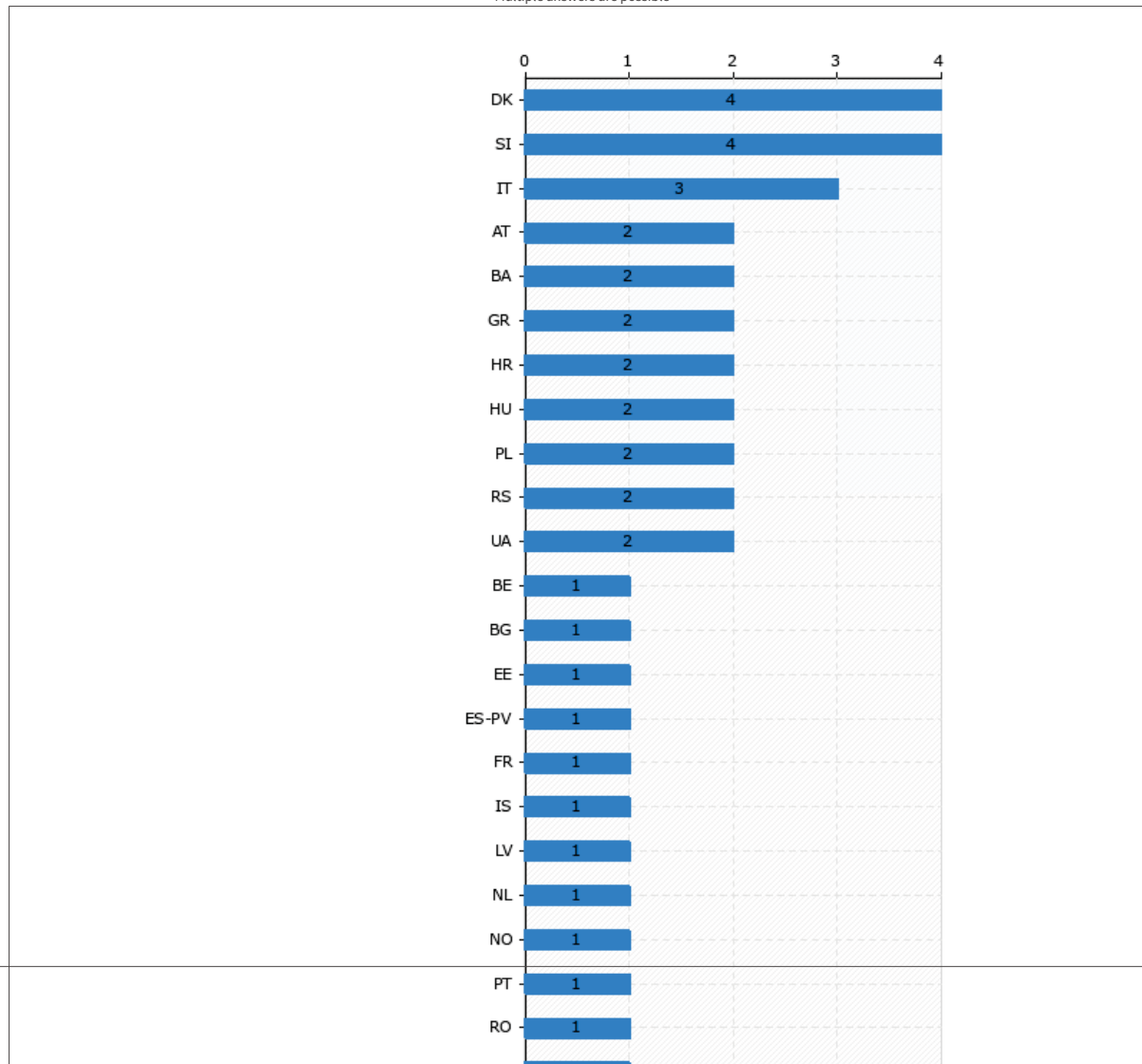# ANALYSIS - CHARTS

**Please select a partner country you represent: (n = 35)**
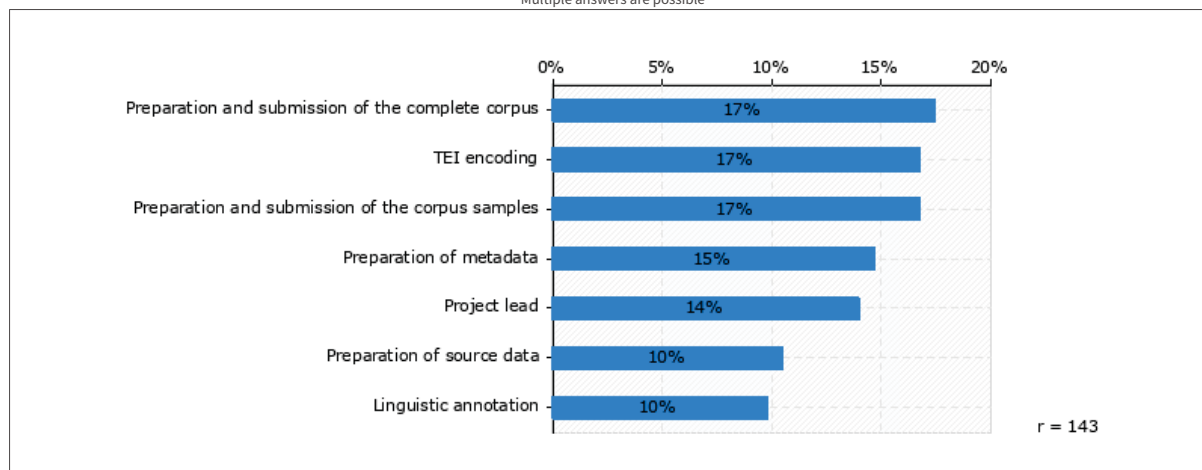
Multiple answers are possible

| Country | Value |
|---------|-------|
| DK | 4 |
| SI | 4 |
| IT | 3 |
| AT | 2 |
| BA | 2 |
| GR | 2 |
| HR | 2 |
| HU | 2 |
| PL | 2 |
| RS | 2 |
| UA | 2 |
| BE | 1 |
| BG | 1 |
| EE | 1 |
| ES-PV | 1 |
| FR | 1 |
| IS | 1 |
| LV | 1 |
| NL | 1 |
| NO | 1 |
| PT | 1 |
| RO | 1 |

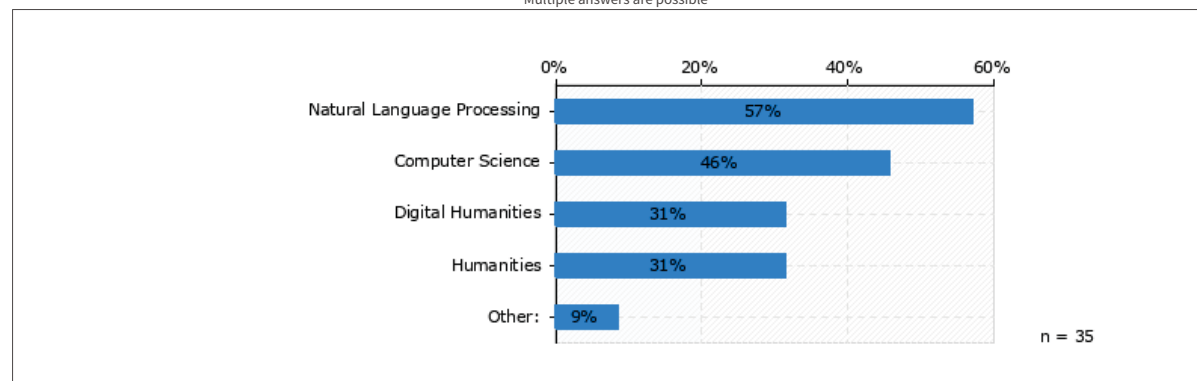Were you a part of the ParlaMint I phase of the project? (n = 35)
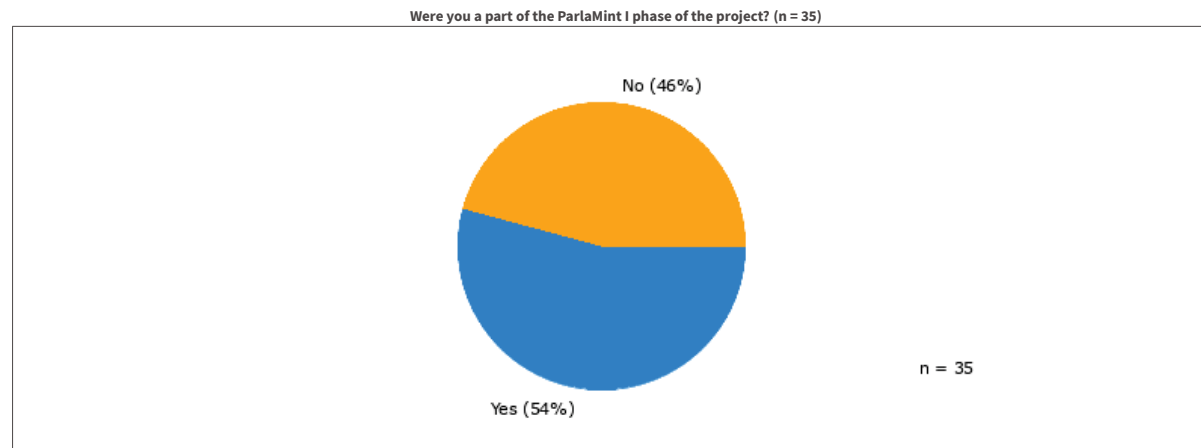
**What was your role in the project workflow? (n = 143)**

Multiple answers are possible

**What is your background? (n = 35)**
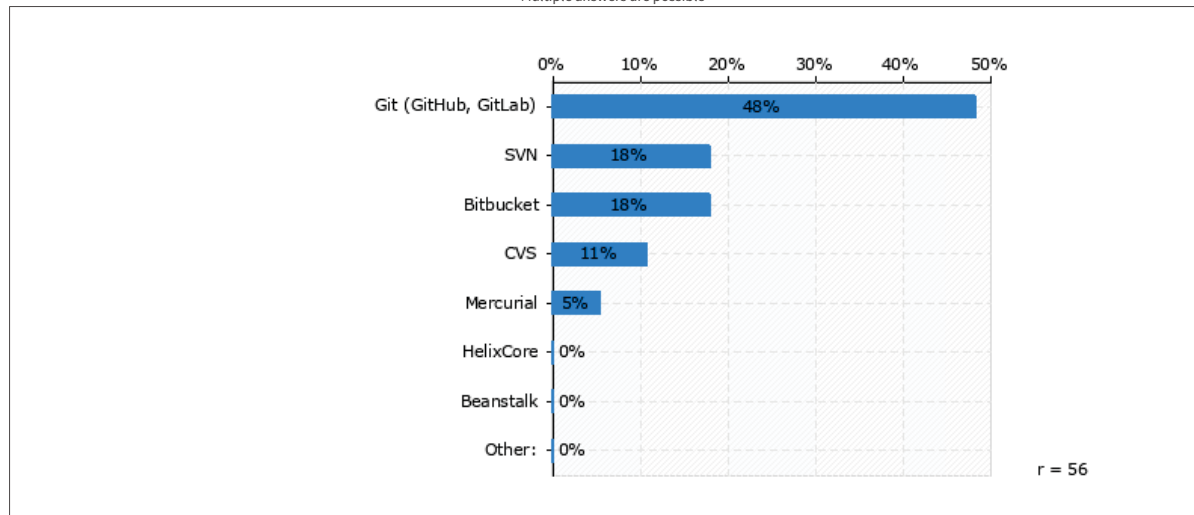
Multiple answers are possible

| Background | Percent |
|---|---|
| Natural Language Processing | 57% |
| Computer Science | 46% |
| Digital Humanities | 31% |
| Humanities | 31% |
| Other: | 9% |

n = 35

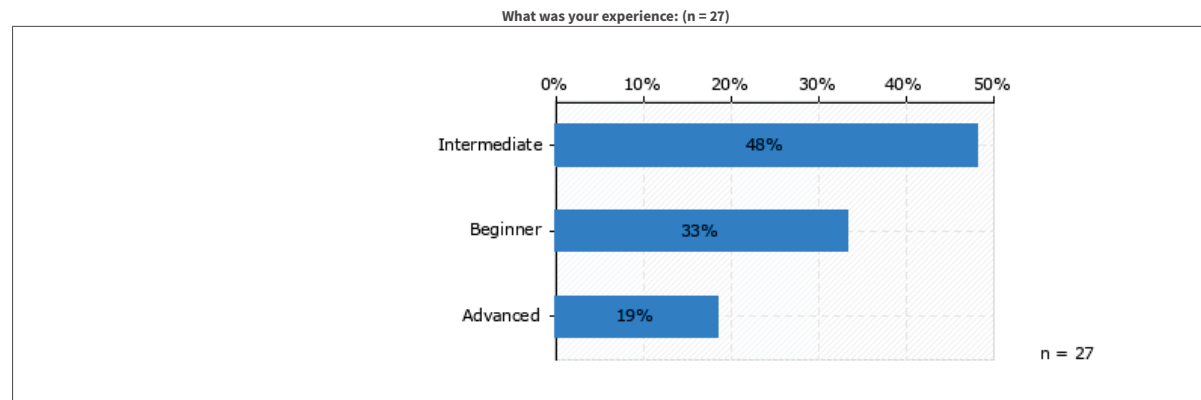| Q4e_text | Q4 (Other: ) | | | | |
|---|---|---|---|---|---|
| | Answers | Frequency | Percent | Valid | Cumulative |
| | linguistics | 1 | 3% | 33% | 33% |
| | machine learning | 1 | 3% | 33% | 67% |
| | physics | 1 | 3% | 33% | 100% |
| Valid | Valid | 3 | 9% | 100% | |
| | -2 (Skipped question (IF logic)) | 32 | 91% | | |
| | Valid | 32 | 91% | | |
| | Total | 35 | 100% | | |

**Do you have previous experience with version control systems like Git? (n = 35)**
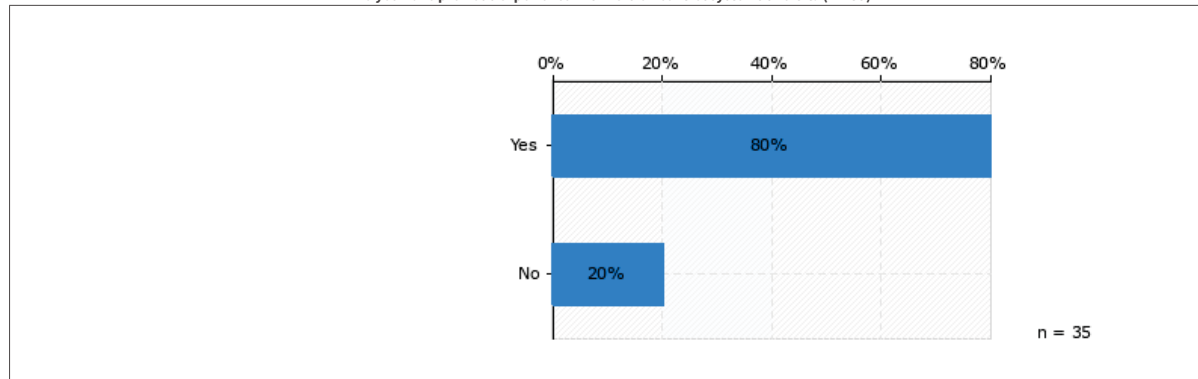
**Which VCS have you used? (n = 56)**

Multiple answers are possible

| VCS | Percentage |
|-----|-----------|
| Git (GitHub, GitLab) | 48% |
| SVN | 18% |
| Bitbucket | 18% |
| CVS | 11% |
| Mercurial | 5% |
| HelixCore | 0% |
| Beanstalk | 0% |
| Other: | 0% |

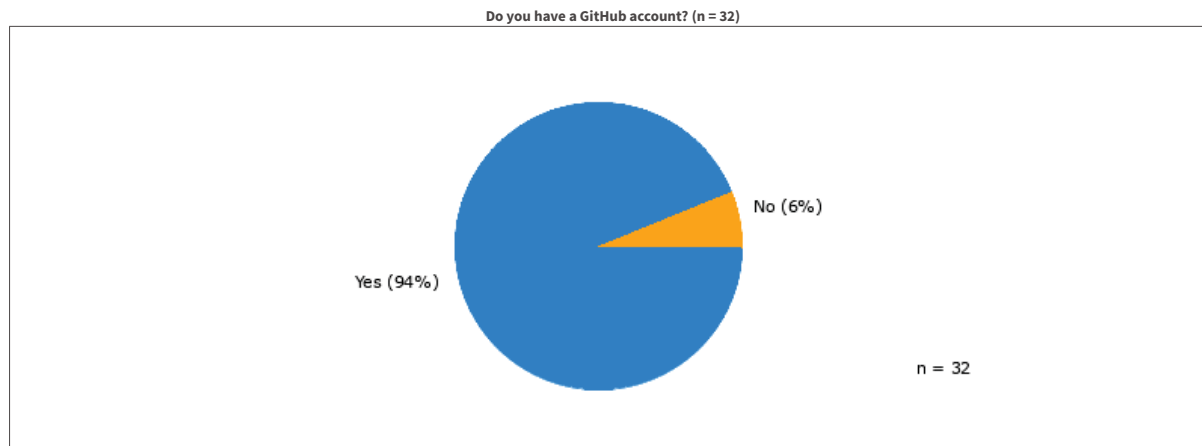r = 56

**What was your experience: (n = 27)**

**Which OS do you use for your work with Git? (n = 33)**

Multiple answers are possible

**Do you have a GitHub account? (n = 32)**



Yes (94%)

No (6%)

n = 32

**Have you used Git as part of ParlaMint I+ II projects? (n = 34)**



No (15%)

Yes (85%)

n = 34
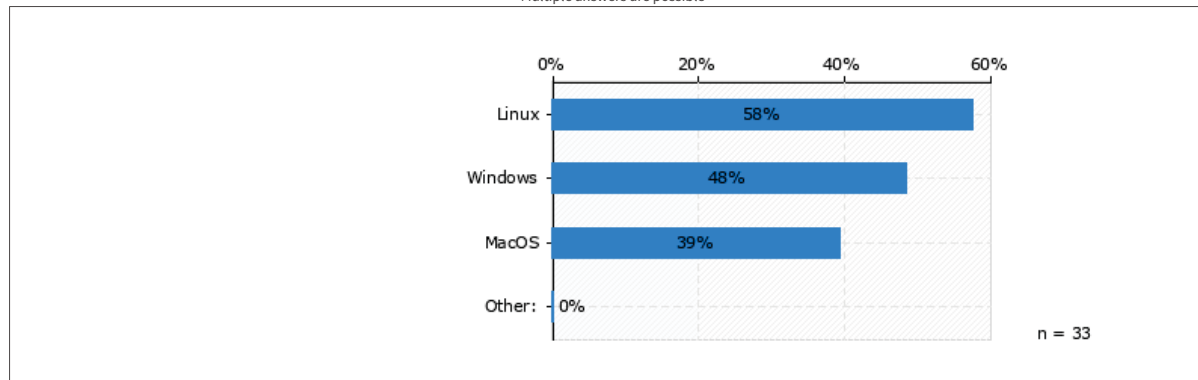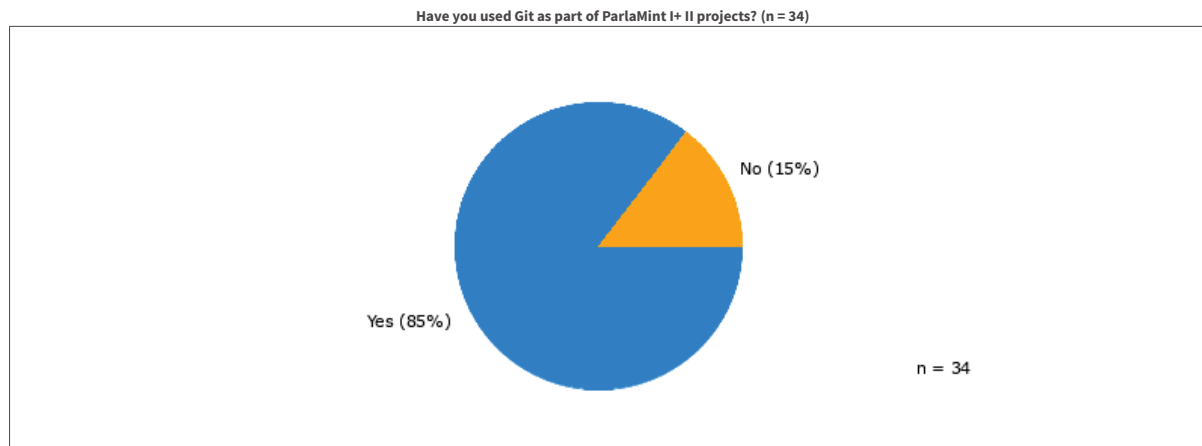
Did you create a GitHub issue to discuss any issues you encountered? (n = 34)



No (32%)

Yes (68%)

n = 34
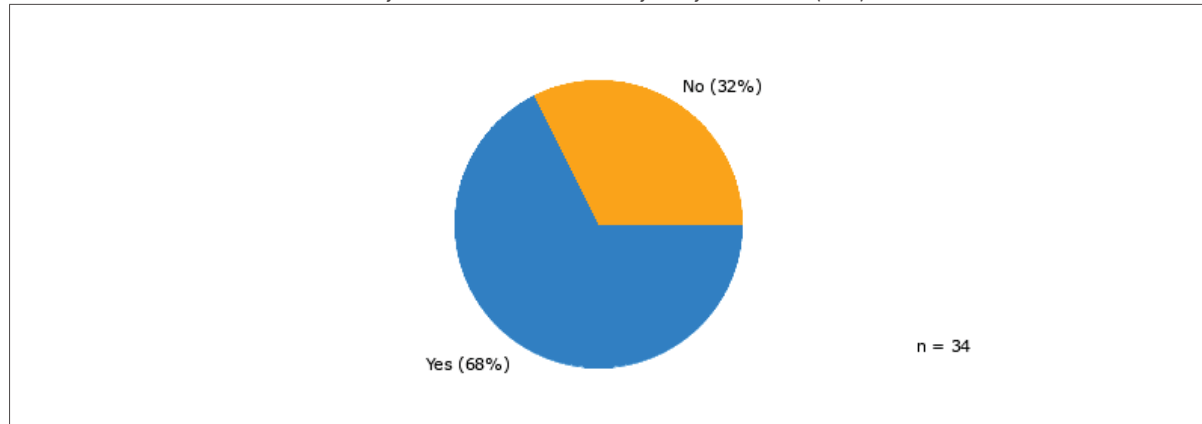
Did you use GitHub Issues to search for information about your issue before you published it? (n = 23)



Yes (96%)

No (4%)

n = 23
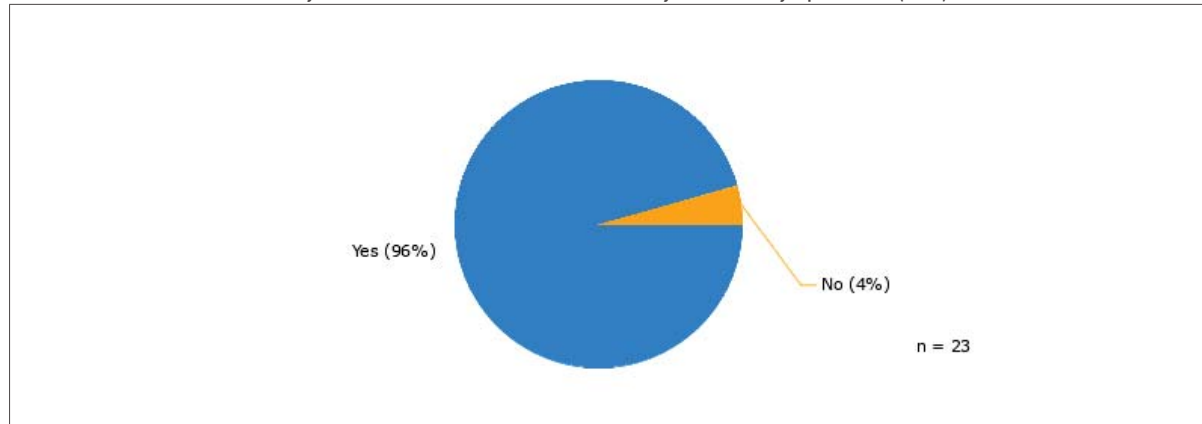
Were you able to get the information you were looking for and solve your problem described in the issue? (n = 23)



Yes (96%)

No (4%)

n = 23
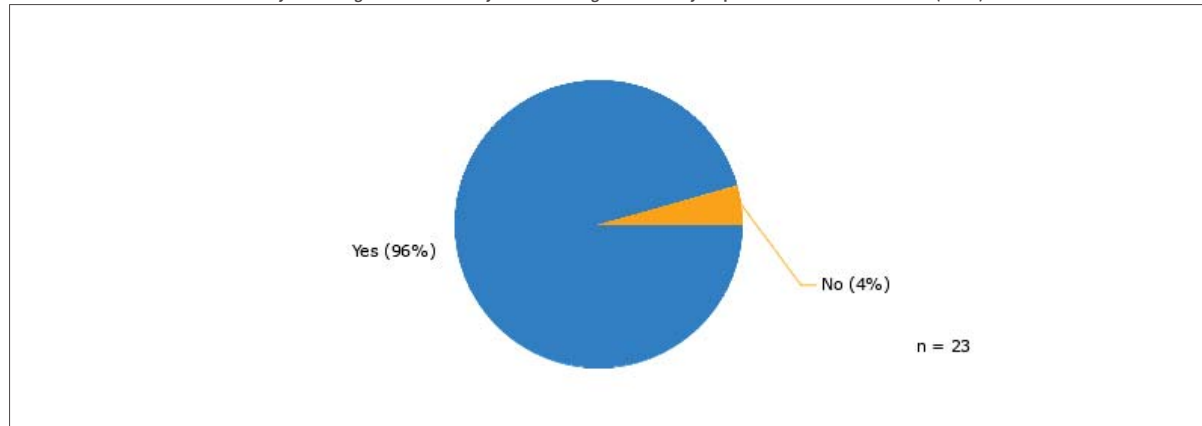
**Did you respond to comments in the Issue you posted (or any other issue)? (n = 33)**



No (24%)

Yes (76%)

n = 33
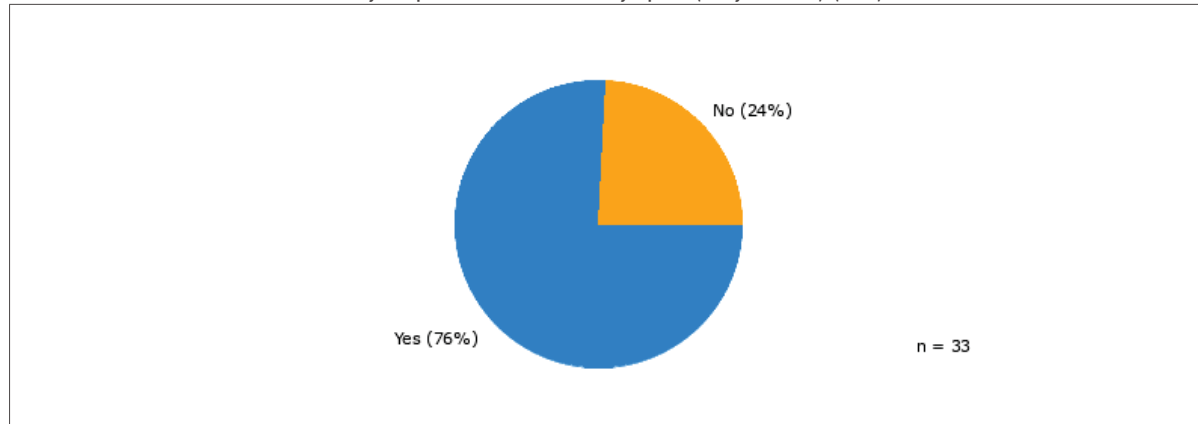
**Were the Git requirements and workflow clearly explained in the instructions (README and CONTRIBUTING files)? (n = 30)**



No (23%)

Yes (77%)

n = 30
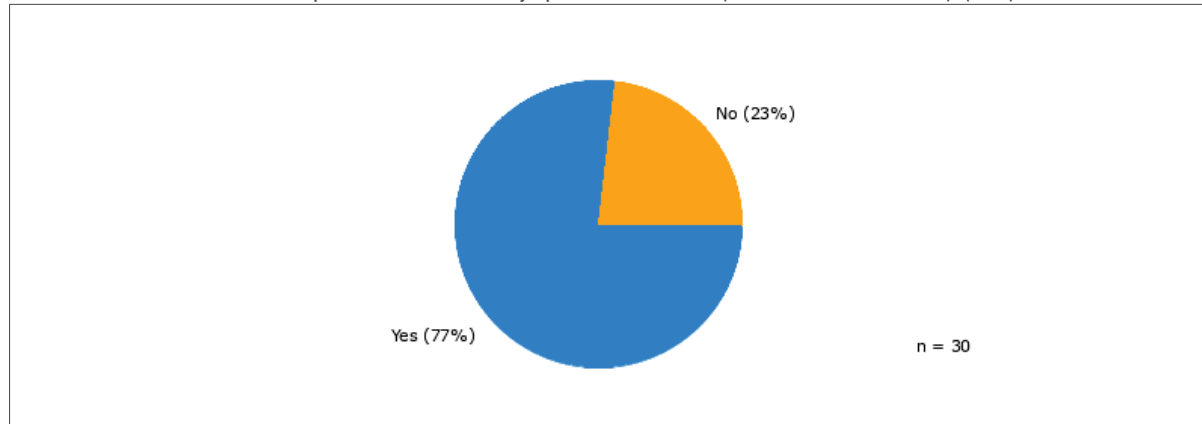
**Please describe what was unclear:**

| Q17 | Please describe what was unclear: | | | | |
|---|---|---|---|---|---|
| | Answers | Frequency | Percent | Valid | Cumulative |
| | - | 1 | 3% | 14% | 14% |
| | it is in the issues | 1 | 3% | 14% | 29% |
| | some instructions in the contributing file should have been in the readme file | 1 | 3% | 14% | 43% |
| | some tei information regarding error messages were not clear. | 1 | 3% | 14% | 57% |
| | readme linked to parla-clarin instead of parlamint, upper limit of sample size was not clearly stated, paths to java libraries were hardcoded, some of the tests assume linux (vs osx), unclear that connlu and vertana were autogenerated | 1 | 3% | 14% | 71% |
| | the whole process with getting the sample approved first was unclear to me | 1 | 3% | 14% | 86% |
| | the branches. the cases when pull request was needed and the cases where no pull request was needed - but these may be obvious for others. | 1 | 3% | 14% | 100% |
| Valid | Valid | 7 | 20% | 100% | |
| | -2 (Skipped question (IF logic)) | 24 | 69% | | |
| | -3 (Drop-out) | 4 | 11% | | |
| | Valid | 28 | 80% | | |
| | Total | 35 | 100% | | |

Did you experience any particular difficulties in submitting data samples to GitHub? (n = 30)

Yes (30%)

No (70%)

n = 30
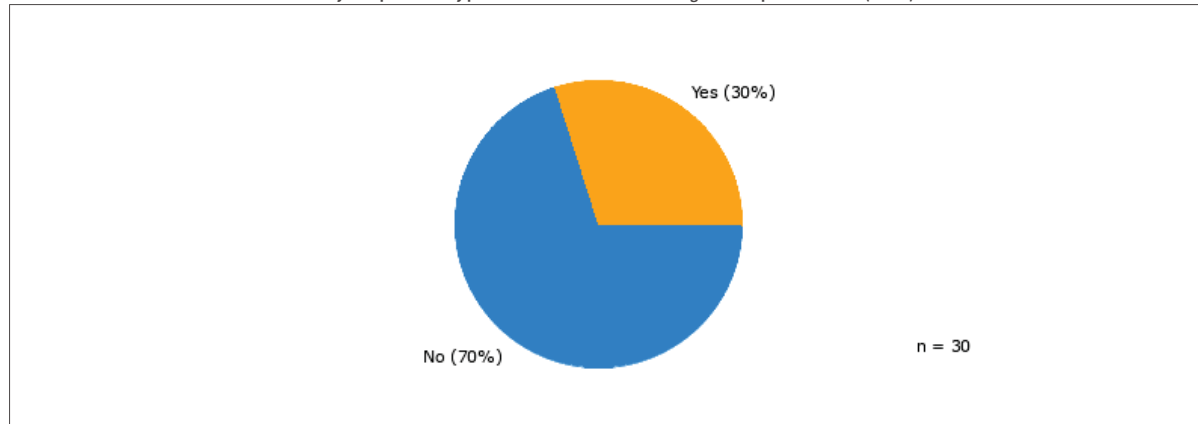
**Please describe in more detail what the difficulties were:**

| Q19 | Please describe in more detail what the difficulties were: | | | | |
|---|---|---|---|---|---|
| | Answers | Frequency | Percent | Valid | Cumulative |
| | my first sample was too big. i did not know github would run validation automatically on push. | 1 | 3% | 11% | 11% |
| | i had no idea about the size restrictions so i produced an initial sample which had to be downsized | 1 | 3% | 11% | 22% |
| | repository forks and synchronization | 1 | 3% | 11% | 33% |
| | the size limit of the data sample was not clear for me before first submitting the sample. and, again, to pull request or not to pull request... :) | 1 | 3% | 11% | 44% |
| | we have support from parlamint, but the conversation was not very fluent, regarding some errors we had in the process. | 1 | 3% | 11% | 56% |
| | i would not call it \ "difficulties\ "-proper: it was not clear, that i should submit a sample (say 3) of \ "full\ " tei files and you would strip down these to the files to a absolute minimum (only a few utterances) yourself. that wasted a bit of effort on my side. | 1 | 3% | 11% | 67% |
| | errors in the parlamint-tei - several rounds of review | 1 | 3% | 11% | 78% |
| | due to my inexperience the automatic merge did not work because my repository was not up to date with clarin-eric/parlamint | 1 | 3% | 11% | 89% |

| | | | | | |
|---|---|---|---|---|---|
| | differentiating between the sample and full delivery (eg. [parlamint] v [parlamint sample]) made for a convoluted workflow in terms of version control | 1 | 3% | 11% | 100% |
| Valid | Valid | 9 | 26% | 100% | |
| | -2 (Skipped question (IF logic)) | 22 | 63% | | |
| | -3 (Drop-out) | 4 | 11% | | |
| | Valid | 26 | 74% | | |
| | Total | 35 | 100% | | |

**Did you post your issues to the GitHub Issues? (n = 9)**



No (33%)

Yes (67%)

n = 9
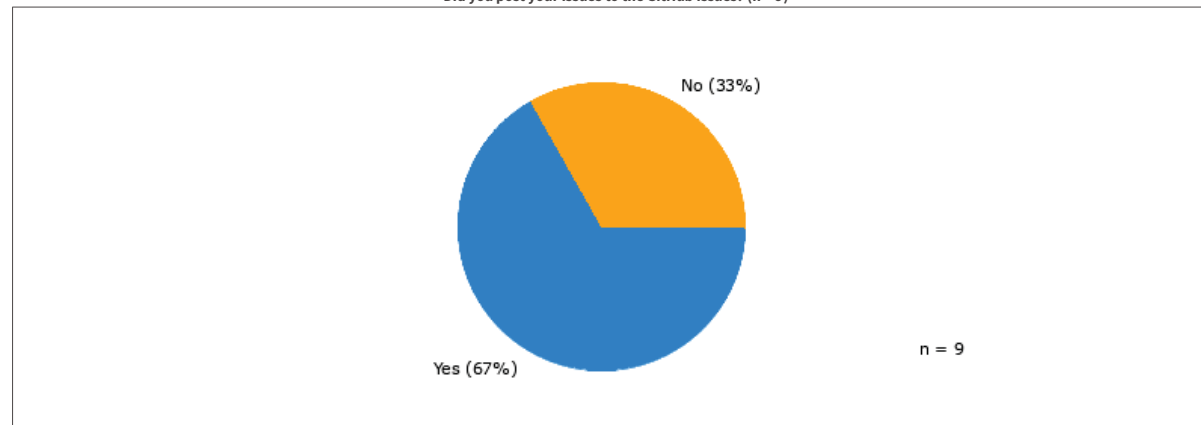
**Were you able to get sufficient support from the project team? (n = 6)**
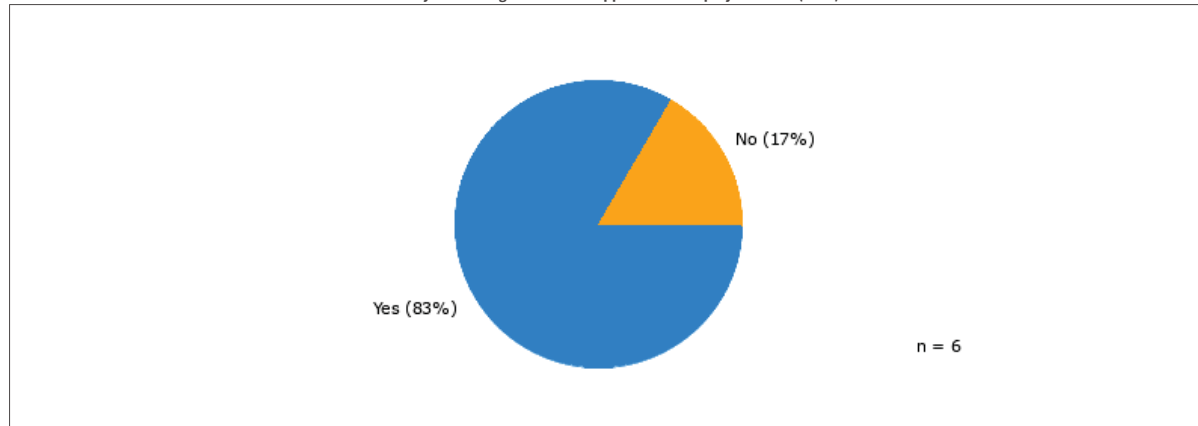


No (17%)

Yes (83%)

n = 6

**Were there any other features provided by the GitHub platform that you think eased communication and the work process on the project?**

| Q22 | Were there any other features provided by the GitHub platform that you think eased communication and the work process on the project? | | | | |
|---|---|---|---|---|---|
| | Answers | Frequency | Percent | Valid | Cumulative |
| | other participants\ ' issues did answer quite a few questions we also had. | 1 | 3% | 17% | 17% |
| | validation of the corpus | 1 | 3% | 17% | 33% |
| | diffs and links to specific lines of a specific file of a specific commit. | 1 | 3% | 17% | 50% |
| | yes surely the general features helped | 1 | 3% | 17% | 67% |
| | it was the 1st time that i was using github issues as a kind of communication platform. i think the one important feature to make it a success wasn\ 't the \ "github platform\ " but the people (tomaz! matyas!) behind it, who were incredibly responsive to all issues at hand! | 1 | 3% | 17% | 83% |
| | i think the possibility to make pull requests and submit samples via automatics validation really streamlined the sample delivery process | 1 | 3% | 17% | 100% |
| Valid | Valid | 6 | 17% | 100% | |
| | -1 (Unanswered question) | 22 | 63% | | |
| | -3 (Drop-out) | 7 | 20% | | |
| | Valid | 29 | 83% | | |
| | Total | 35 | 100% | | |

**Git workflow and Issues (n = 26)**



| | |
|---|---|
| I received relevant information/feedback after posting an issue | 4.4 |
| I plan to use Git in my future work | 4.3 |
| GitHub Issues eased communication and the work process on the project | 4.3 |
| GitHub Issues proved to be an effective means of communicating with editors and other project partners. | 4.2 |
| The acquired knowledge about Git will be useful for my future work | 4.2 |
| I was able to clearly communicate issues/ideas using Git Issues | 4 |
| The Git requirements and workflow were clearly explained in the instructions (README and CONTRIBUTING files) | 3.8 |
| I had no particular difficulty in submitting the data sample | 3.3 |

n = 26

1 - Strongly disagree    2 - Somewhat disagree    3 - Neutral    4 - Somewhat agree    5 - Strongly agree

**Any additional feedback related to workflow and submission process?**

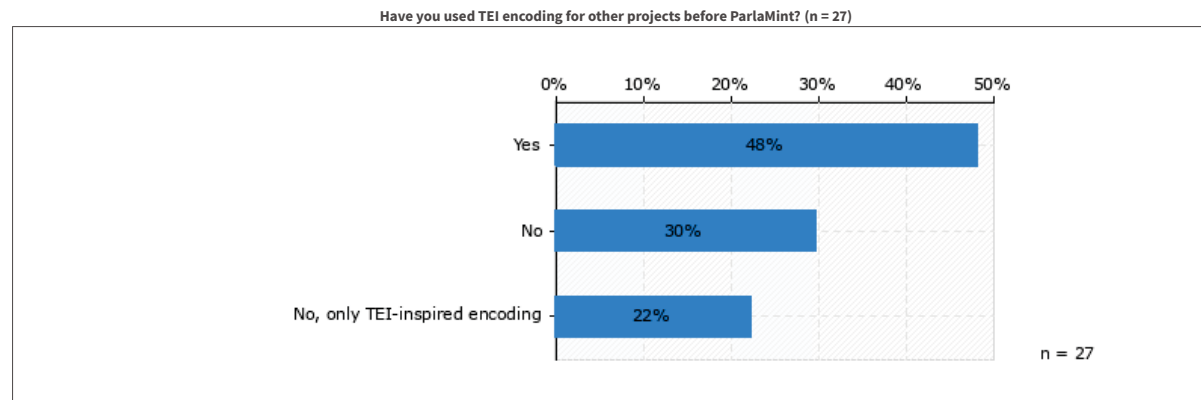| Q24 | Any additional feedback related to workflow and submission process? | | | | |
|---|---|---|---|---|---|
| | Answers | Frequency | Percent | Valid | Cumulative |
| | i helped preparing the samples/the data, but a colleague submitted it in github. we had some problems related to the fact that new versions of the material in github were uploaded without being notified. | 1 | 3% | 17% | 17% |
| | parlamint issues and git issues are different topics | 1 | 3% | 17% | 33% |
| | the existence of the github issues really helped me in troubleshooting my own problems / bugs as i could search for other countries with similar problems, which made the troubleshooting much easier | 1 | 3% | 17% | 50% |
| | video tutorial / workshop | 1 | 3% | 17% | 67% |
| | i cannot stress enough how heplful people have been every time i posted an issue. they were able to perfectly understand what the issue was, which is sometimes difficult to put into words, and then reply in an extremely useful way | 1 | 3% | 17% | 83% |
| | i suggest running an xml formatter on commit, so that every commit is formatted in the same way. this would facilitate diffing and ensure consistency across corpora. | 1 | 3% | 17% | 100% |
| Valid | Valid | 6 | 17% | 100% | |

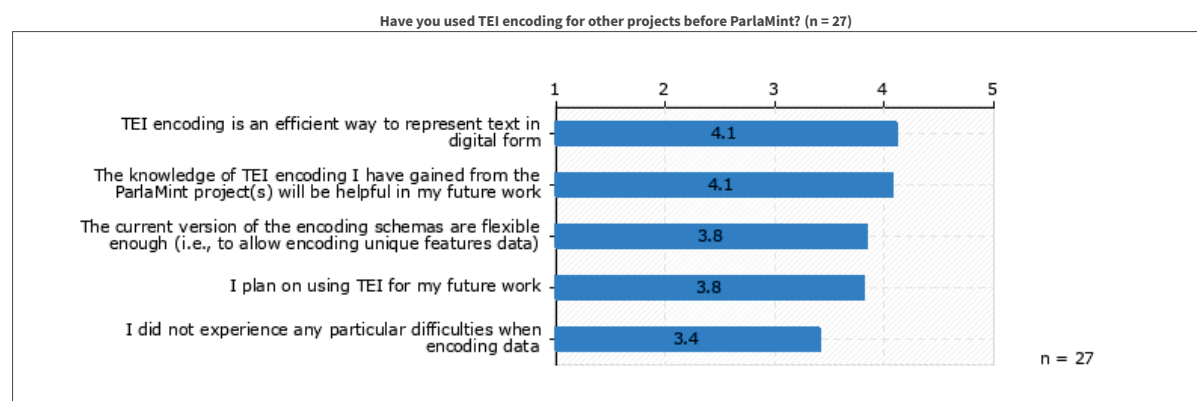| | | | | |
|---|---|---|---|---|
| | -1 (Unanswered question) | 22 | 63% | | |
| | -3 (Drop-out) | 7 | 20% | | |
| | Valid | 29 | 83% | | |
| | Total | 35 | 100% | | |

**How familiar were you with the TEI P5 Guidelines or TEI in general before you started working on ParlaMint? (n = 27)**
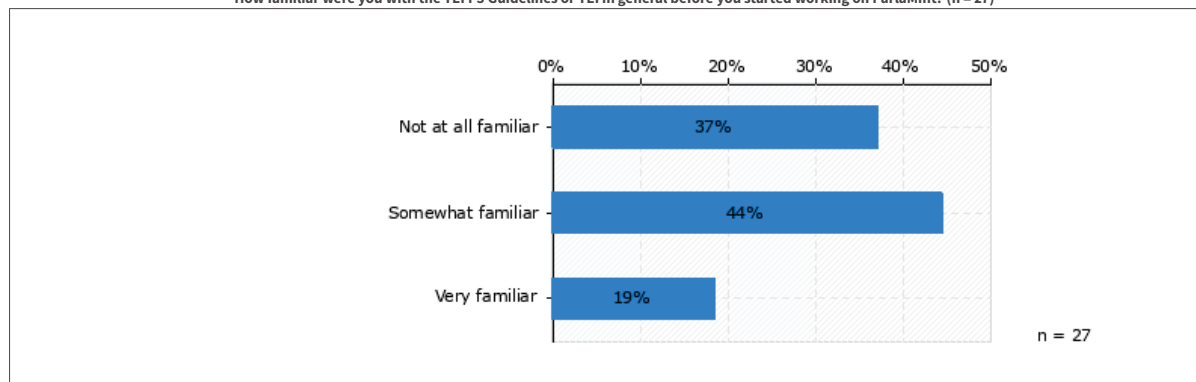
Did you need additional help (external sources) to become familiar with TEI before you started working on the ParlaMint project(s)? (n = 22)



Yes (36%)

No (64%)

n = 22

**Have you used TEI encoding for other projects before ParlaMint? (n = 27)**

Have you used TEI encoding for other projects before ParlaMint? (n = 27)
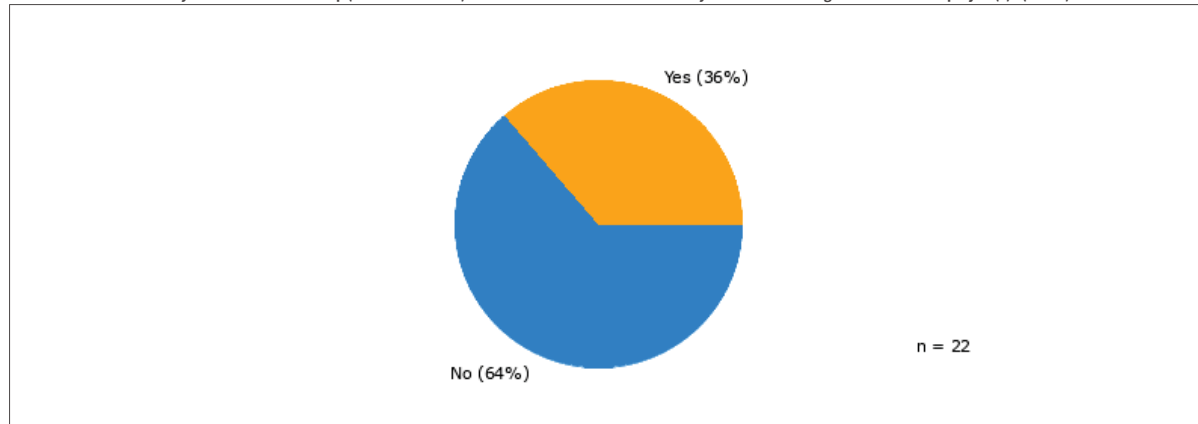
**Any additional feedback regarding the TEI encoding?**

| Q30 | Any additional feedback regarding the TEI encoding? | | | | |
|---|---|---|---|---|---|
| | Answers | Frequency | Percent | Valid | Cumulative |
| | minimize \ "optional\ " features that end up being not so optional. | 1 | 3% | 33% | 33% |
| | not sure that a significant part of the parlamint community was aware of the complexity of the tei encoding process, me included. this should be communicated in the future (is not easy to assess, i know). | 1 | 3% | 33% | 67% |
| | i believe my difficulty using tei came mostly from my little knowledge about the tei norm and about my own parliament, along with the difficulty to express parliamentary data, as well as trying to make it uniform through very different parlimentary systems and languages. all that put together made it a nightmare, but i don\ 't think it could be made much easier. we just had to work … | 1 | 3% | 33% | 100% |
| Valid | Valid | 3 | 9% | 100% | |
| | -1 (Unanswered question) | 24 | 69% | | |
| | -3 (Drop-out) | 8 | 23% | | |
| | Valid | 32 | 91% | | |
| | Total | 35 | 100% | | |