



CEvNS data sharing

Janina Hakenmüller, prepared for the discussion session

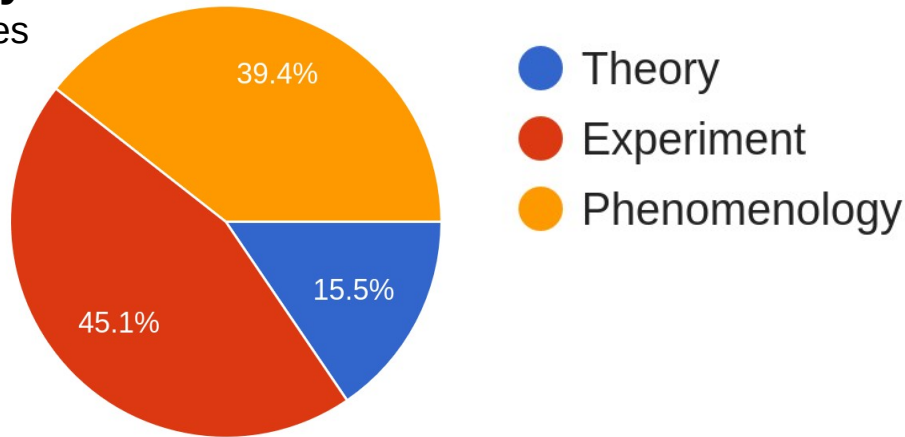


Participants

52 responses in total → Many thanks to everybody!

What is your main focus?

70 replies



Collaborations:

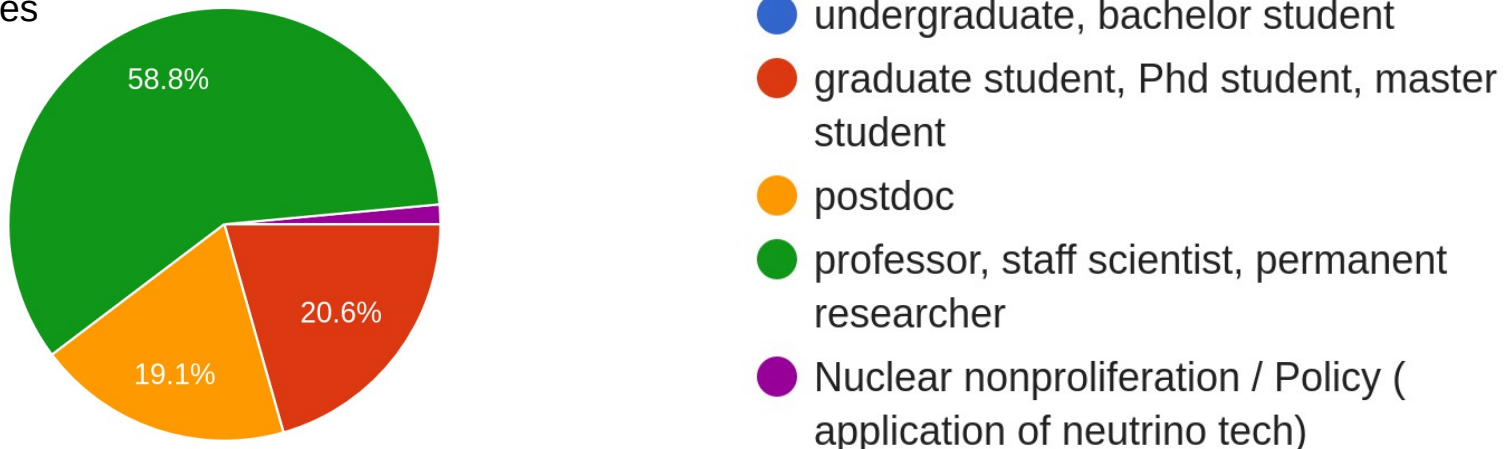
(voluntary information)

- | | |
|---------------|------------|
| COHERENT (10) | CCM (2) |
| NUCLEUS (3) | RELIS |
| CONNIE | CDEX |
| TEXONO | LEGEND (2) |
| CONUS (2) | PROSPECT |
| RED-100 | DarkSide |
| NuGEN | |
| Miner | |
| SBC | |
| PALEOCCENE | |
| CRAB | |
| nuESS | |

10 spokespersons

What is your current position?

67 replies



Requirements and platforms

Are you required from funding agencies to make your data public?

Experiments (collaborations):

47% (mostly US, Europe)

53%

(remark: do not necessarily need to make the raw data public, but need data management plan in experiment and theory)

Theory, phenomenology:

yes

11%

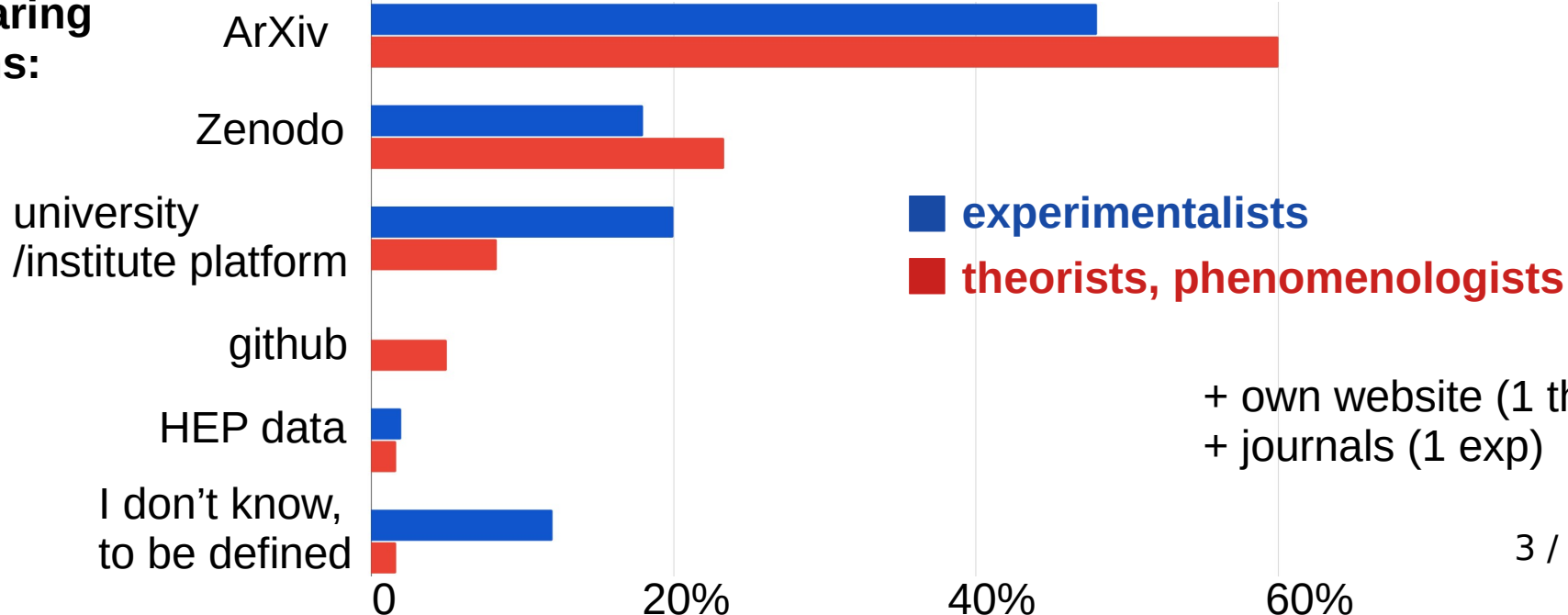
no

32%

not applicable

(mostly phenomenologists) 57%

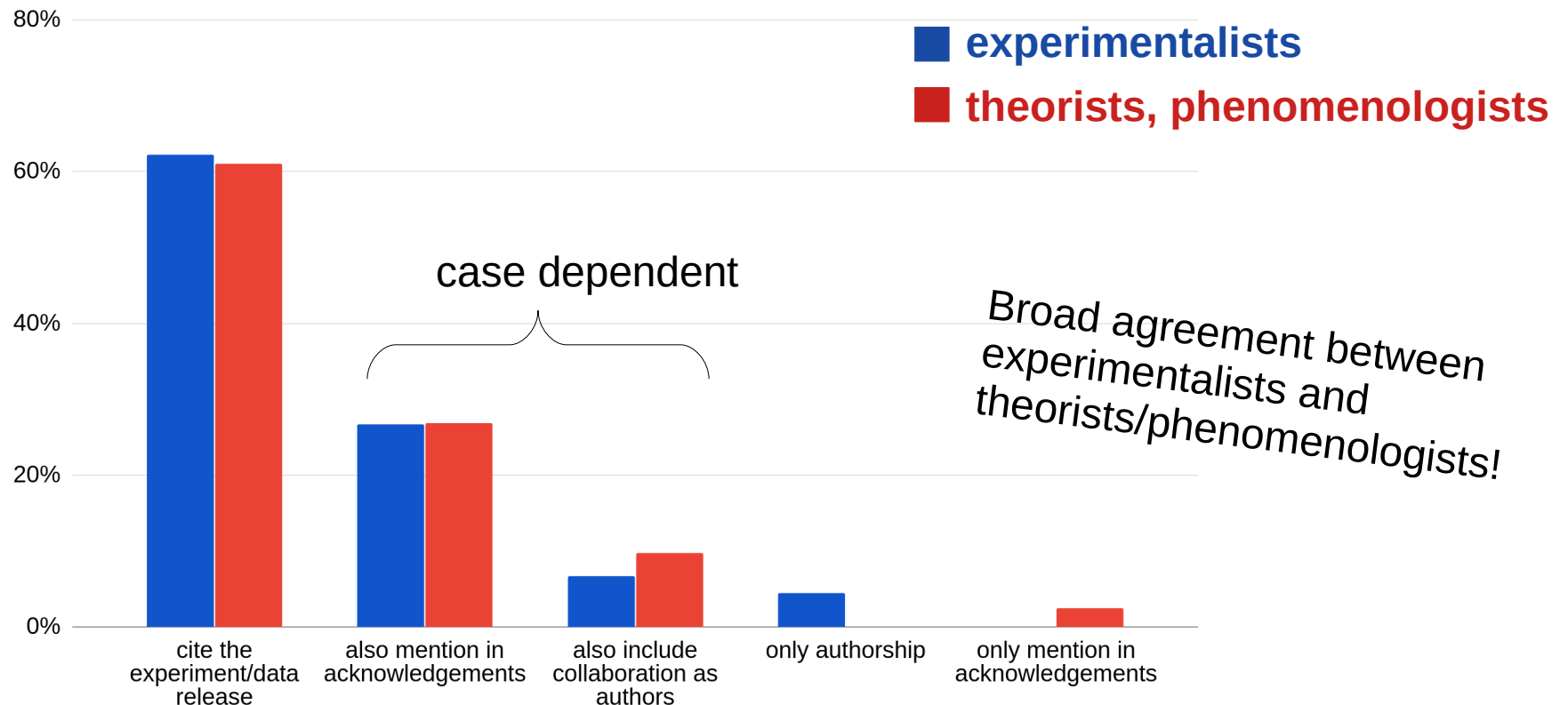
Data sharing platforms:



+ own website (1 theo)
+ journals (1 exp)

Credit to experiments

If public data is used in a publication, how do you think the experiments/experimenters should be credited?

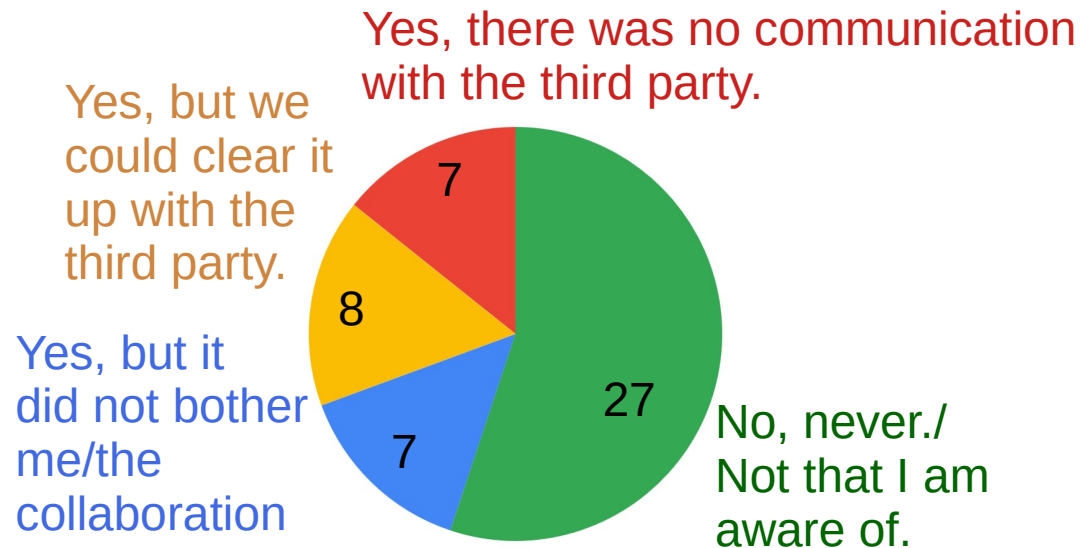


Spokepersons from two experiments:

- write explicitly in the text in the context how the data was used
- If the collaboration helped in the use of data, acknowledgment or even authorship is warranted

Issues and concerns

Did you ever have the impression that your data/information provided in a publication were used in an **incorrect way by a third party**?



(response "not applicable" mostly by pheno not shown)

Have you ever been **discouraged** by an experimentalist from **working with their data**?

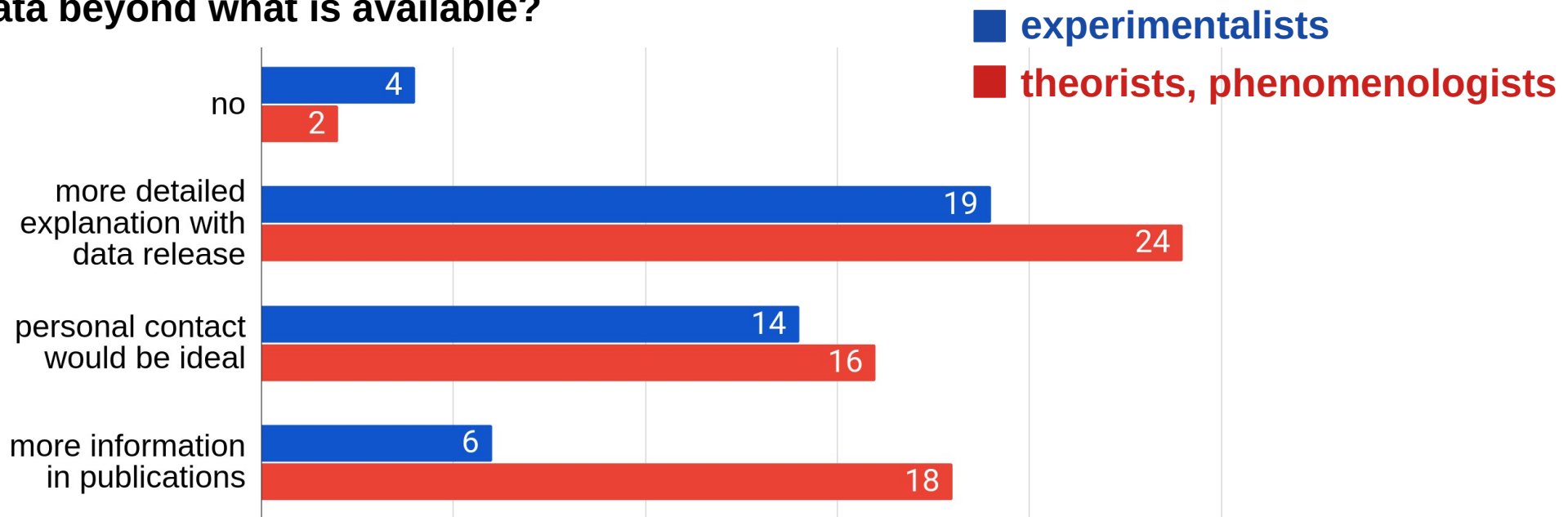
- Experimentalists: **3 yes**, 16 no, 12 not applicable
- Theorists/Phenomenologists: **8 yes**, 24 no, 7 not applicable

Direct statements by Phenomenologists:

- outside the CEvNS community broadly data is usually not available
- no replies to emails
- Data release did not match figures in the publication
- Given private figures for efficiencies

Improvements, ideas & suggestions

When working with third party data, do you think you need more information on the data beyond what is available?



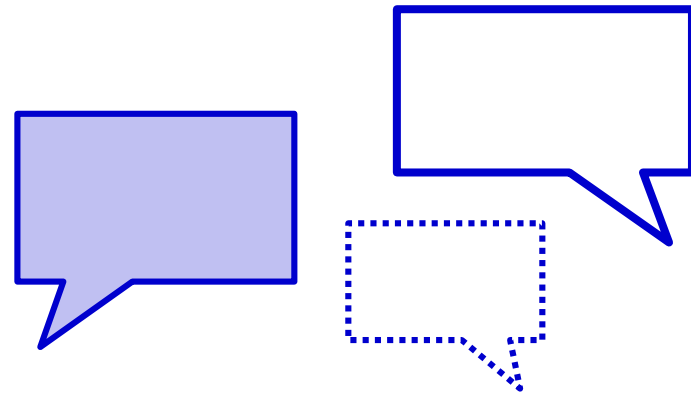
Data format:

- common way to present data between experiments/standardized form of data releases
- keep the format for the data release similar from one measurement to the next to avoid needed to rewrite scripts

Additional information people are interested in:

- small sample of code would be ideal as with the 2017 Coherent Csl data release
- peer reviewed data sets to ensure not information is missing?
- unpublished "private communication" information, have explicit explanations in text
- README to reproduce plots in publications

Discussion



Ideas:

- [standardized form](#) of data release, common presentation? (several comments)
- [peer-reviewed](#) data releases?
- [super-repository](#) that links all data releases?
- more collaborations (neutrino and dark matter communities) should have data releases (2 comments)
- [feedback loop](#) with theory community on what to include, how useful the release was
- ...

Risks/concerns:

- It is not obvious that the [efforts--return ratio](#) is favorable to have (construct and maintain) a public archive site of CEvNS data.
- the obvious "danger" is that the [data are used incorrectly](#) (curated data release might not provide access to the fully detailed treatment of systematics, but event level data cannot be interpreted correctly without help of original data collector)
- I don't like to work with third party data due to the [risk of misunderstandings](#)
- the [danger of being "scooped"](#) on searches with one's own data, prevent especially for thesis
- ...