PsychOpen CAMA

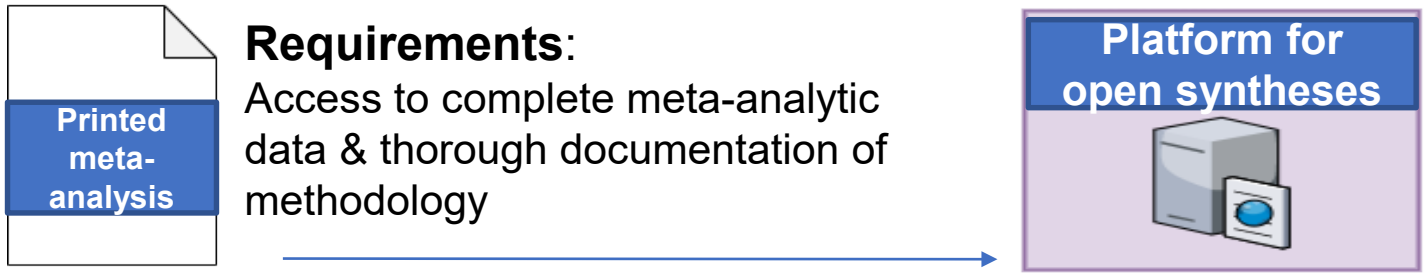
A platform for open and cumulative meta-analyses in psychology

ESMA-R Conference, January, 22nd, 2021

Tanja Burgard, ZPID – Leibniz Institute for Psychology

Contact: tb@leibniz-psychology.org

Why do we need open syntheses?

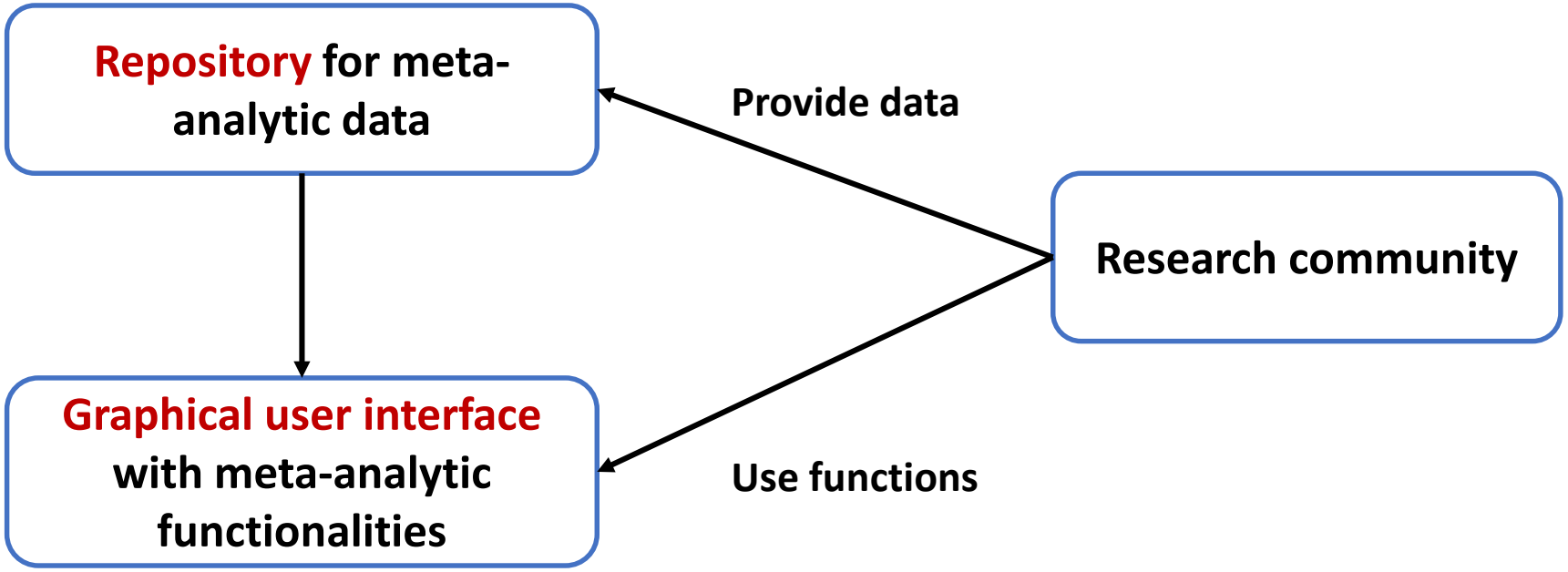


- **Fixed**: Interpretation and discussion of results under certain methodological assumptions and decisions
- **Static**: Overview on a research question at a certain point in time

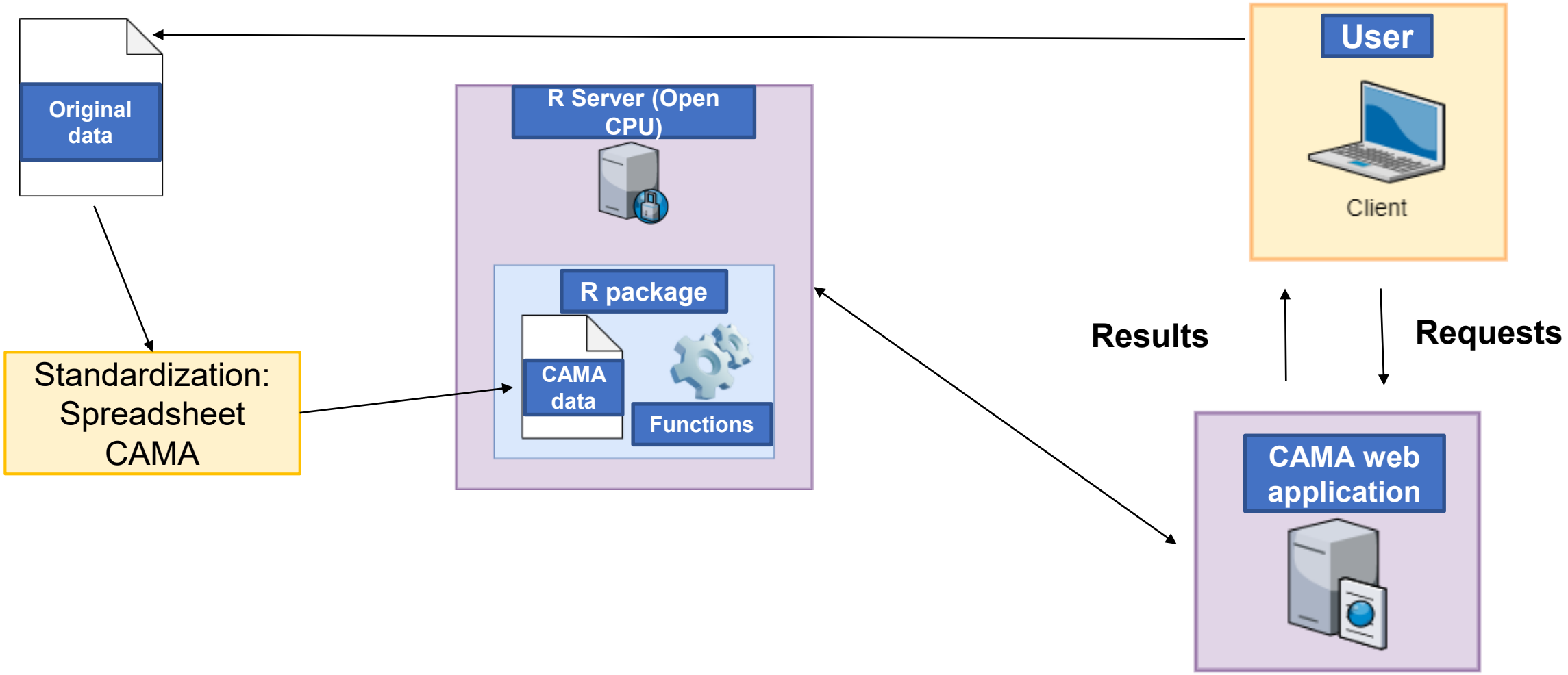
- **Flexible**
  - Analyses can be replicated
  - Robustness of results can be checked by varying subjective decisions
- **Dynamic**
  - Data can be re-used for other analysis purposes
  - Data can be updated more efficiently and faster

The basic concept of CAMA

CAMA = Community Augmented Meta-Analysis



The architecture of PsychOpen CAMA



Data Inspection

select domain	Personality psychology
Select dataset	Sex differences in Narcissism

Original publication	Prinz, G. M. (2019). Sex differences and trait intercorrelations in the dark triad of personality: A systematic review and psychometric meta-analysis [Unpublished master's thesis], University of Vienna.
Data provider	Lisa Bucher
License	CC BY 4.0
Research question	Univariate and multivariate sex differences in subclinical components of the Dark Triad of personality
Inclusion and exclusion criteria	Inclusion: (1) Studies need to investigate the three components of the Dark Triad (narcissism, Machiavellianism, psychopathy) within one sample and (2) report results (mean and standard deviation) of men and women for the three traits
Effect size	Hedge's g for every component, Mahalanobis' D as a multivariate measureIndividual corrections for measurement error variances will be applied for Hedge's g, following recommendations and formulas of Wiernik and Dahlke (2020), use of R package 'psychmeta'
Hierarchical structure	If a study reported several effect sizes for a sample, only one effect size was culled to ensure statistical independence of samples in the synthesis

Filter

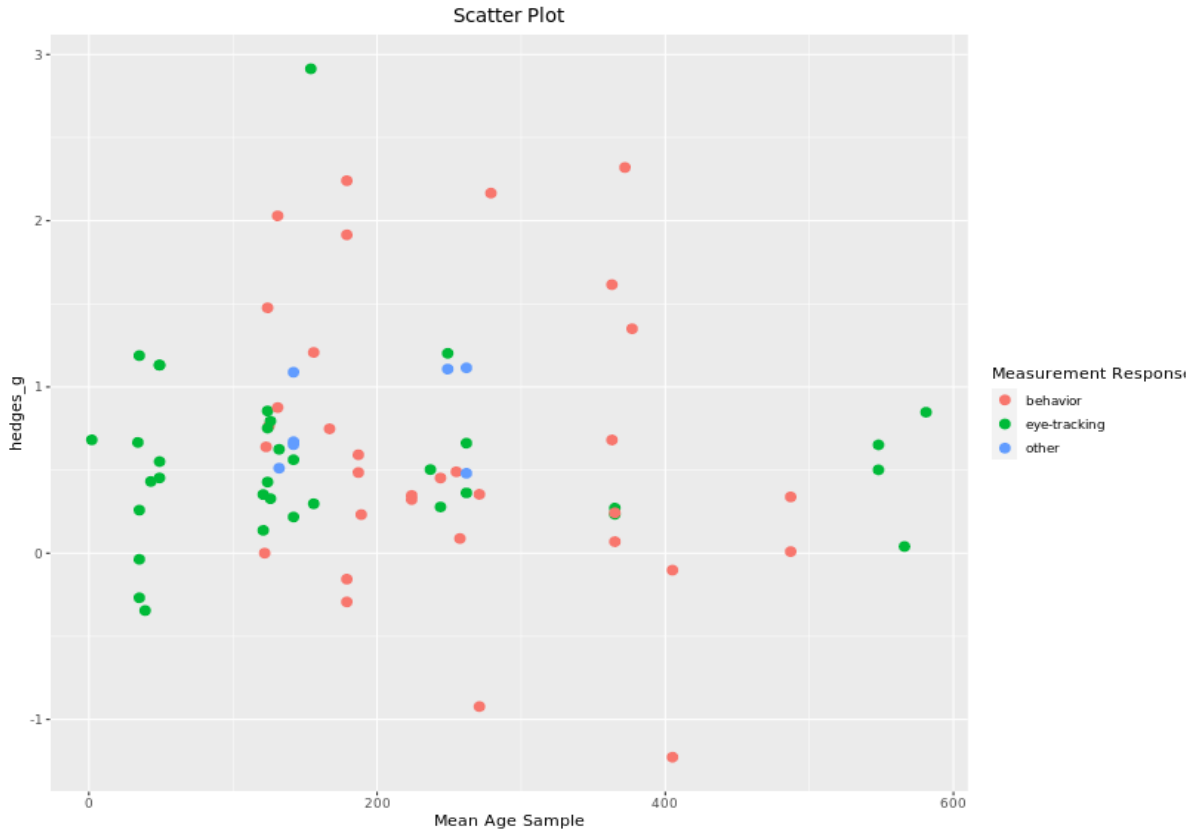
report_ID	r_author	r_year	r_peer	r_cites	r_past	r_citesav	r_lab	r_language	s_targetpop	s_n	s_nmale	s_nfemale	s_meanage	s_sdage	s_country
Adler (2017)	Adler	2017	no	2	3	0.6667	other	english	mixed	499	185	314	22	3.7	USA
Aghababaei et al. (2014)	Aghababaei	2014	yes	67	6	11.1667	other	english	adult	223	90	133	31.24	8.94	Iran
Anderson & Cheers (2017)	Anderson	2017	yes	19	3	6.3333	other	english	adult	173	45	128	23.37	7.88	Australia
Anisdahl Jonson & Venbakken Sagerud	Anisdahl	2017	no	0	3	0	other	english	mixed	159	78	81			Norway

Data Exploration

select domain	Cognitive development
select dataset	Infant directed speech preference
select effect size type	hedges_g

Select Moderators:

- [x] Measurement Response
- [x] Mean Age Sample
- [ ] Speaker



The Scatterplot for a numeric and a categorical variable depicts the relationship between the numeric moderator and the effect size of the meta-analysis. Moreover, different groups of the categorical moderator are grouped by colour.

Basic Analyses

select domain	Personality psychology
select dataset	Correlation of Narcissism and Psychopathy
select effect size type	pearson_r

Select Moderators:

- [ ] Mean Age Sample
- [x] Publication Year
- [ ] Publication Status
- [ ] Language Report
- [ ] Target Population
- [ ] Sample Size
- [ ] Sample Size Women
- [ ] Sample Size Men
- [ ] Language Measurements
- [ ] Scale Narcissism
- [ ] Scale Psychopathy
- [ ] Number of Items on Narcissism
- [ ] Number of Items on Psychopathy
- [ ] Number of Items on Dark Triad
- [x] Laboratory
- [ ] Scale Type
- [ ] OECD Status

Update View

RMA Model | Forest Plot | Cumulative Forest Plot

```
Mixed-Effects Model (k = 170; tau^2 estimator: REML)

tau^2 (estimated amount of residual heterogeneity):     0.0275 (SE = 0.0036)
tau (square root of estimated tau^2 value):             0.1657
I^2 (residual heterogeneity / unaccounted variability): 88.47%
H^2 (unaccounted variability / sampling variability):   8.67
R^2 (amount of heterogeneity accounted for):            2.41%

Test for Residual Heterogeneity:
QE(df = 167) = 1411.0954, p-val < .0001

Test of Moderators (coefficients 2:3):
QM(df = 2) = 5.7445, p-val = 0.0566

Model Results:

             estimate      se     zval    pval    ci.lb    ci.ub
intrcpt        0.5368  0.0161  33.3951  <.0001   0.5053   0.5683  ***
r_year        -0.0230  0.0143  -1.6041  0.1087  -0.0510   0.0051
r_labJonason  -0.0706  0.0339  -2.0817  0.0374  -0.1371  -0.0041    *

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

This output shows the results of the Random Effects Meta-Analysis for k effect sizes. K is the number of studies included in the meta-analysis. The information given is interpreted as follows:

Tau^2
Estimated between-study variance, reflects the amount of heterogeneity among the true effect sizes across studies.

Tau
Estimated standard deviation of underlying true effects across studies, can be used to describe the distribution of true effects.

I^2
Variability may occur between studies (true heterogeneity, tau^2) and within studies (sampling error). I^2 is the percentage of the total variability, that is due to true heterogeneity.

H^2
Relative excess in Q over its degrees of freedom. The ratio of the Q statistic to its degrees of freedom is interpreted as a measure of the extent of heterogeneity.

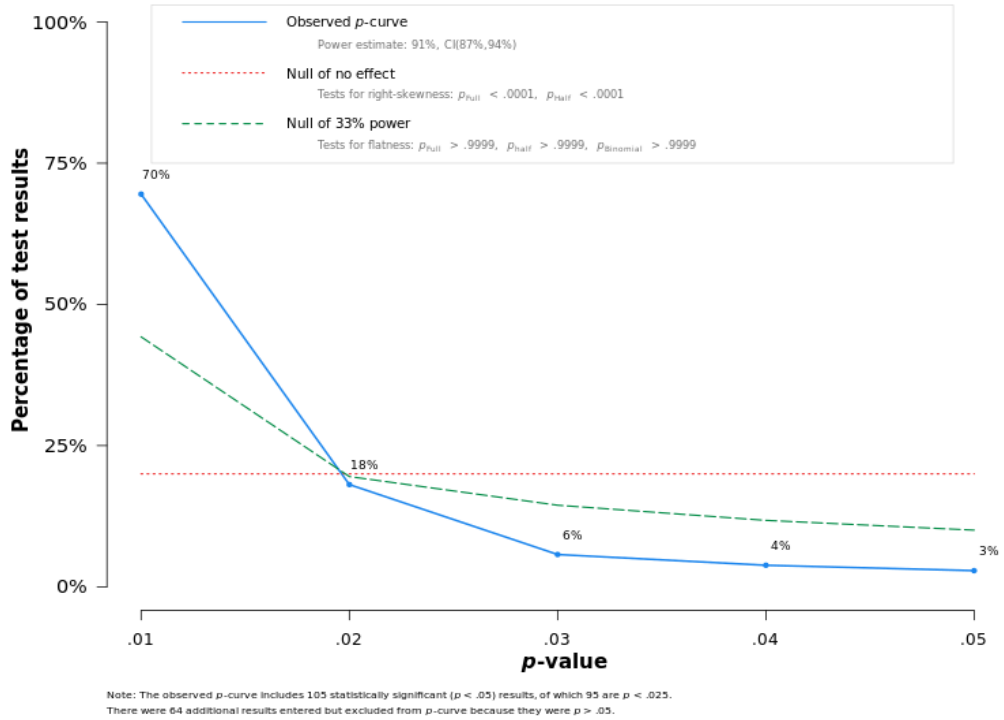
Test for heterogeneity
Cochran Q test for statistical heterogeneity tests the null hypothesis, that the underlying true effect size parameters are the same in all studies included. The test statistic Q is chi-

Publication Bias

select domain	Personality psychology
select dataset	Sex differences in Narcissism
select effect size type	hedges_g_corrected

Update View

Funnel Plot | Contour-enhanced funnel | p-curves



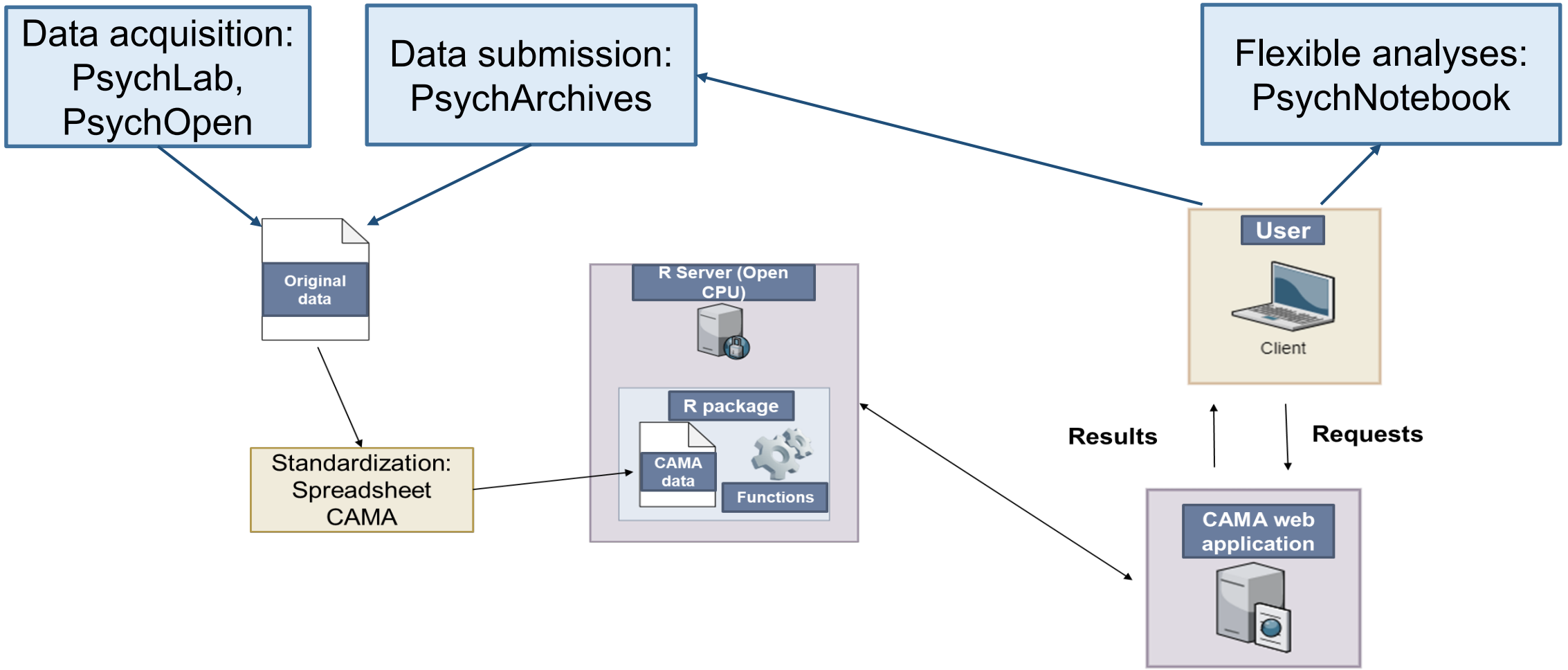
Note: The observed p-curve includes 105 statistically significant ($p < .05$) results, of which 95 are $p < .025$.
There were 64 additional results entered but excluded from p-curve because they were $p > .05$.

p-curves

The p-curve is a tool to assess the evidential value of a set of published findings by examining the distribution of statistically significant p-values (p<0.05). It assumes this distribution to be a function of the real underlying effect. If there is no real effect, the p-values are expected to be uniformly distributed, as the red dotted line in the plot indicates. If there really is an effect, smaller p-values are more likely to be observed, resulting in right-skewed p-curves. The green dashed line shows the shape of a hypothetical curve of 33 % power. In case of higher statistical power of the underlying results, the curve is even more right-skewed. The plot also provides a power estimation for the blue observed p-curve. A left-skewed curve indicates p-hacking, as researchers may stop collecting more data as soon as findings achieve statistical significance. This results in large significant p-values, causing a left-skewed distribution.

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). p-Curve and Effect Size: Correcting for Publication Bias Using Only Significant Results. Perspectives on Psychological Science, 9(6), 666–681. https://doi.org/10.1177/1745691614553988

Synergy effects with other ZPID services



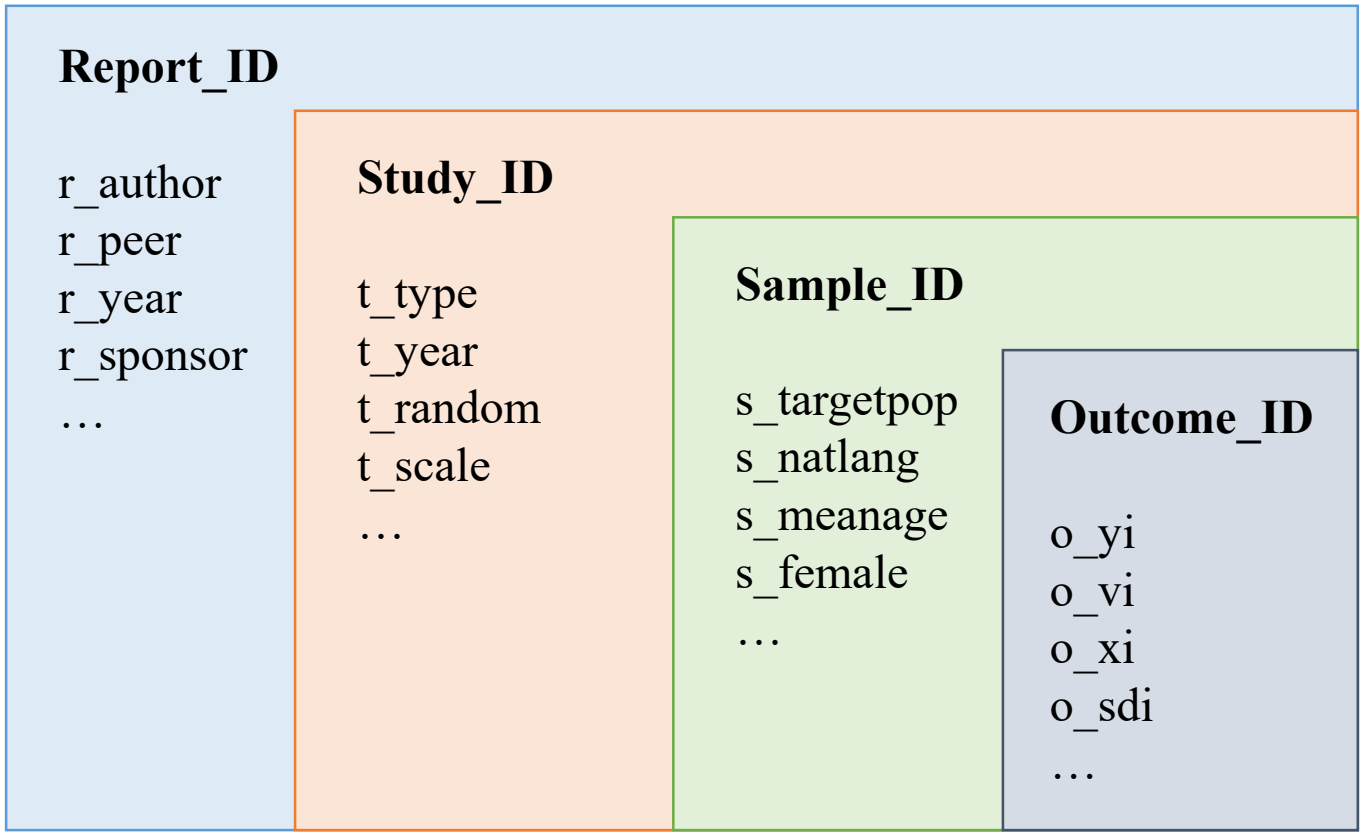
Recommended further reading

- Bastian, H., Glasziou, P., & Chalmers, I. (2010). Seventy-five trials and eleven systematic reviews a day: How will we ever keep up? *PLoS Medicine*, *7*(9), e1000326. https://doi.org/10.1371/journal.pmed.1000326
- Bosco, F., Steel, P., Oswald, F., Uggerslev, K., & Field, J. (2015). Cloud-based Meta-analysis to Bridge Science and Practice: Welcome to metaBUS. *Personnel Assessment and Decisions*, 1(1). https://doi.org/10.25035/pad.2015.002
- Burgard, T., Bosnjak, M., & Studtrucker, R. (in press). Community-Augmented Meta-Analyses (CAMAs) in Psychology. Potentials and Current Systems. *Zeitschrift für Psychologie*.
- Haddaway, N. R. (2018). Open synthesis: On the need for evidence synthesis to embrace open science. *Environmental Evidence*, *7*(1), 4–8. https://doi.org/10.1186/s13750-018-0140-4
- Shojania, K. G., Sampson, M., Ansari, M. T., Ji, J., Doucette, S., & Moher, D. (2007). How quickly do systematic reviews go out of date? A survival analysis. *Annals of Internal Medicine*, *147*(4), 224–233. https://doi.org/10.7326/0003-4819-147-4-200708210-00179
- Tsuji, S., Bergmann, C., & Cristia, A. (2014). Community-augmented meta-analyses: Toward cumulative data assessment. *Perspectives on Psychological Science, 9*(6), 661–665. https://doi.org/10.1177/1745691614552498
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package, *Journal of Statistical Software, 36* (3), 1-48. http://dx.doi.org/10.18637/jss.v036.i03

Further challenges and planned developments

- Implementation of further methodological approaches with different data structures
  - NMA
  - IPD
  - MASEM
- Strategic data acquisition to systematically close research gaps
  - Especially for research questions with a high relevance for decision-making
  - Time-intensive
  - Relies on cooperation of the research community

Data template



Information is collected on different levels. If data are nested, dependencies are automatically accounted for by using a multilevel model

The naming of study outcomes and the type of effect size measure used in the meta-analysis (in the metadata) follows the metafor package (Viechtbauer, 2010). As we use this package as the basis for effect size calculation and most of the meta-analytic outputs, consistent naming makes life easier.