



UK Research
and Innovation



DARE UK

SACRO: Semi-Automated Checking of Research

Outputs

Lizzie Green

Senior Lecturer UWE

Katie Oldfield

Public Engagement Manager

Five Safes and SACRO

Safe
People

Safe
Projects

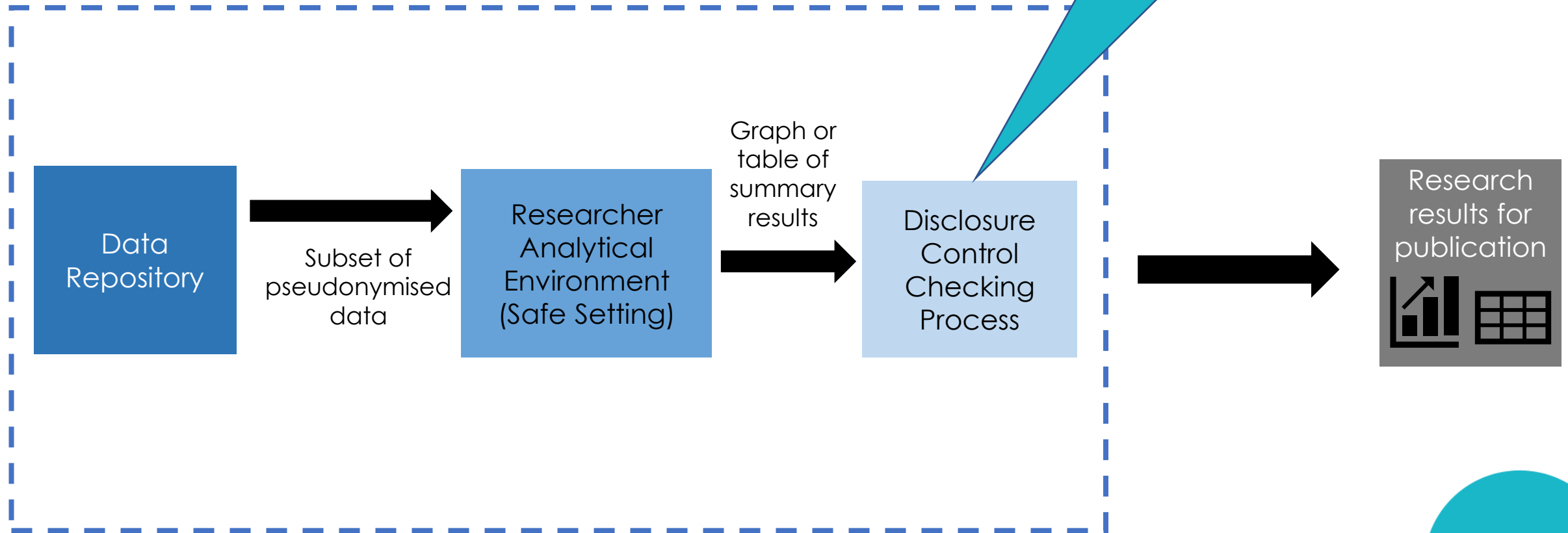
Safe
Setting

Safe
Data

Safe
Outputs



What is the problem?



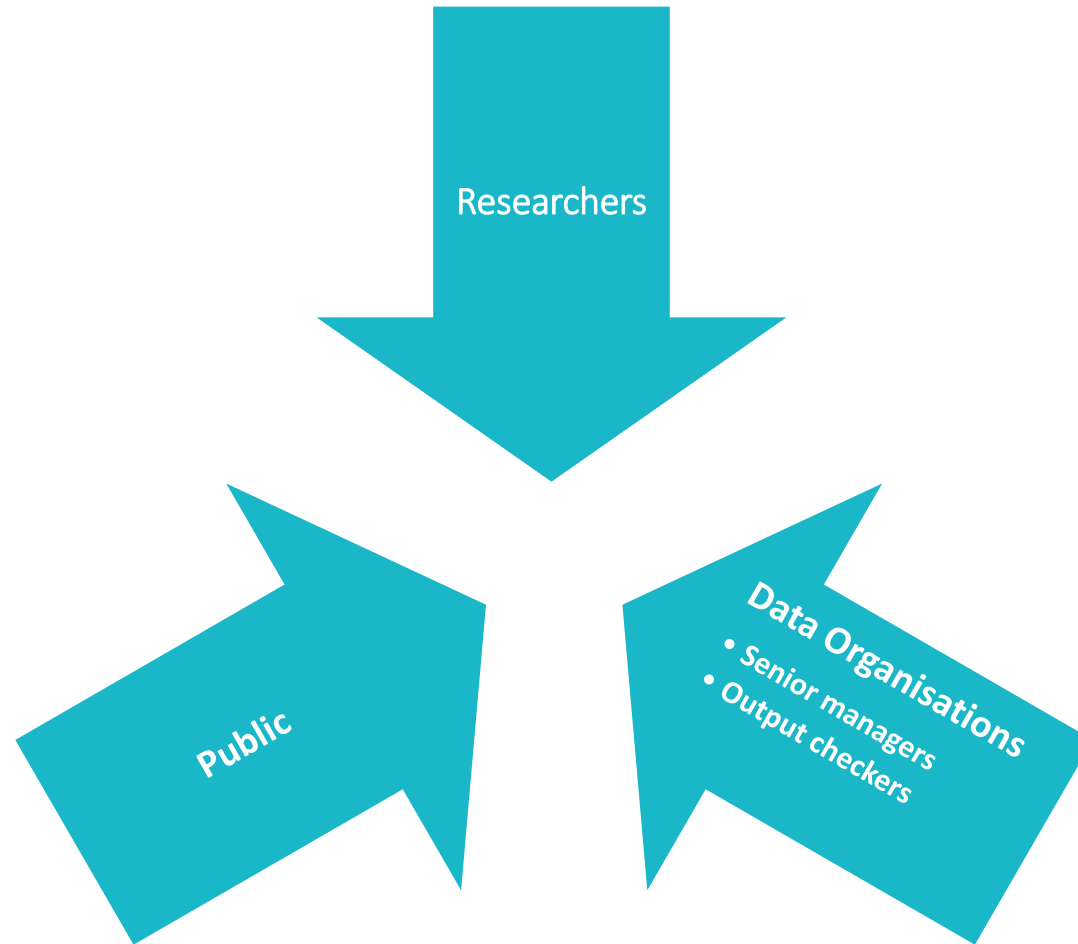
What did we do?

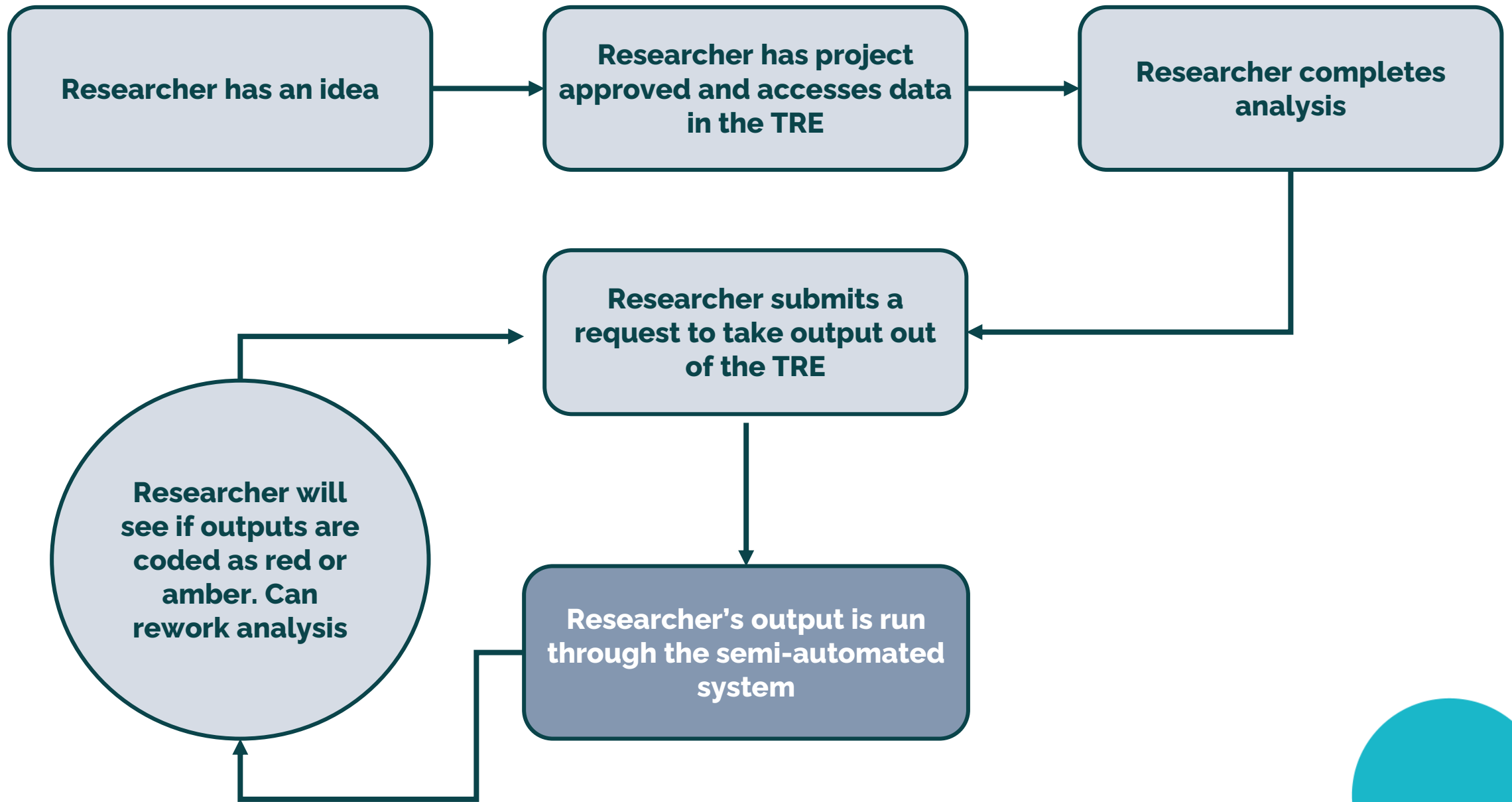
- Design and implement a semi-automated system for checks on common research outputs
 - Immediate feedback to researchers
 - Centralised system for output checkers to look at all the outputs together with a traffic lights system
- Update and modernise guidance to TREs and output checkers to improve consistency
- Consensus statement

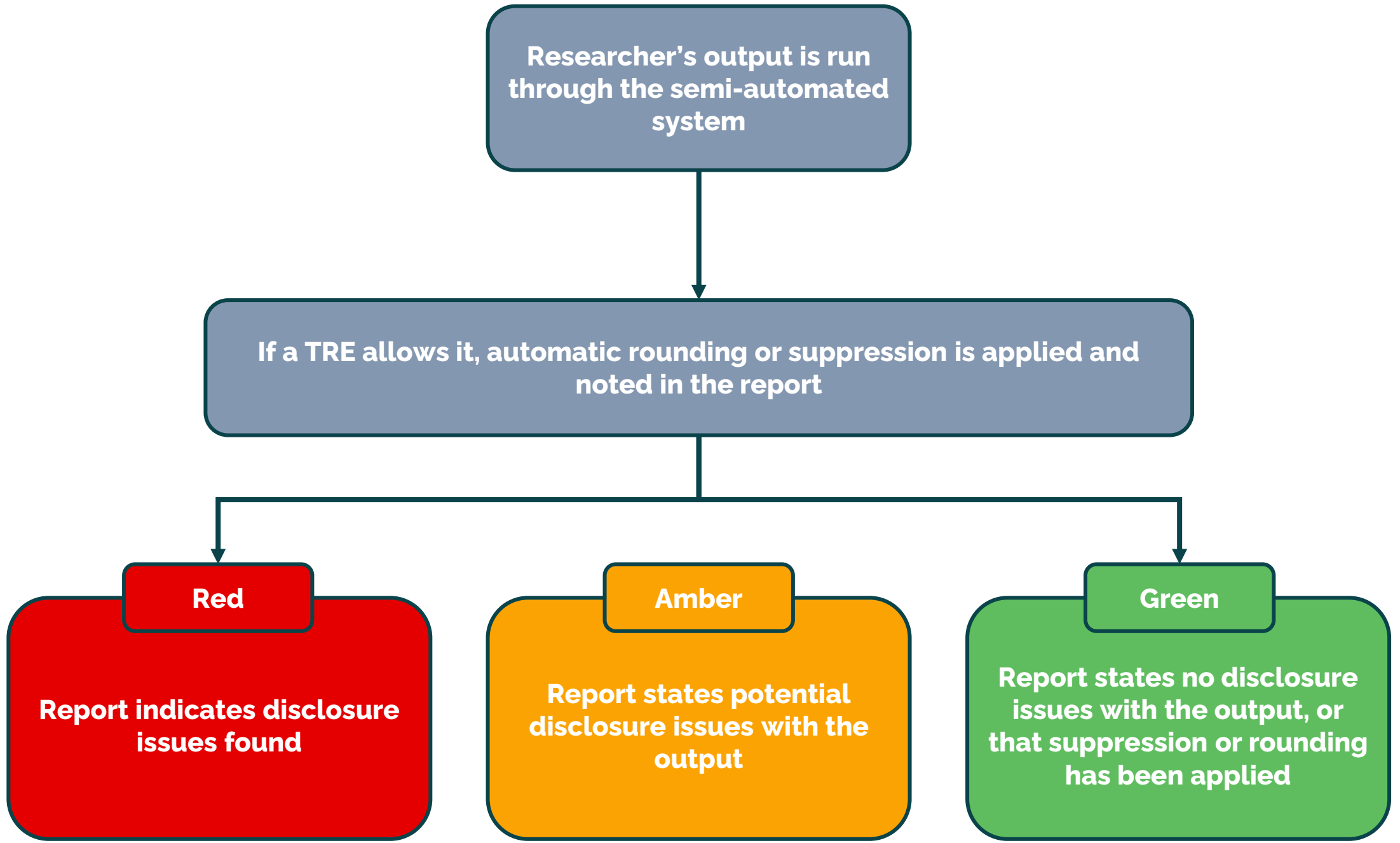


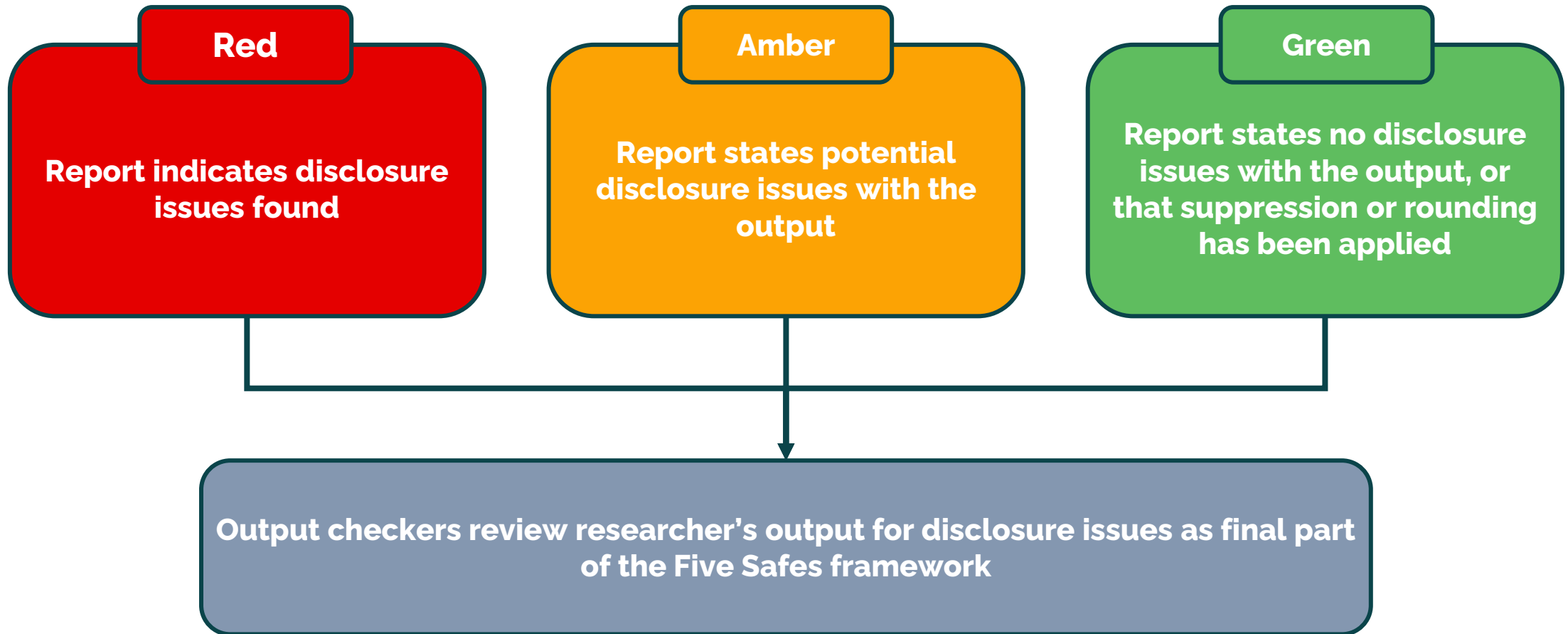


Who did we talk to?









SACRO viewer for output checkers

SACRO Outputs Viewer Release and download

Outputs to review

- output_1 crosstab table
- output_3 pivot_table table
- output_4 ols regression
- output_5 olsr regression
- output_6 probit regression
- output_7 logit regression
- output_8 custom

ACRO risk profile

output_1

Created at: 21 August 2023 at 12:16

Output type: crosstab table

ACRO status: **Fail** (threshold: 4 cells may need suppressing)

Comments:

- Please let me have this data.
- 6 cells were suppressed in this table

Exception request: because I really want it

Review: You cannot approve this output until you add a comment.

Review comments on output_1:

output_1_0.csv Open file

| recommend | great_pret | pretentious | usual |
|------------|------------|-------------|-------|
| not_recom | 1440 | 1440 | 1440 |
| priority | 858 | 1484 | 1924 |
| recommend | 0 | 0 | 2 |
| spec_prior | 2022 | 1264 | 758 |
| very_recom | 0 | 132 | 196 |

Outputs to review

Output with problematic cells highlighted

Comments

Consensus Statement on the use of Automation when Checking the Disclosure Risk of Research Outputs

October 2023



- Explanation about what SACRO does and how it works
- Agreement about how to implement SACRO
- Organisations signed up to it

What does this mean for the future?

- Semi-automating means it frees up output checkers time to focus on complex and nuanced outputs
- Researchers receive immediate feedback from the SACRO tool before sending their requests to output checkers
- Output checkers are working more consistently and to a framework to shared guidance

Reducing the bottleneck means more projects are able to be processed resulting in more researchers being able to carry out their work quicker.



Thank you

Find out more:

sacro.contact@uwe.ac.uk

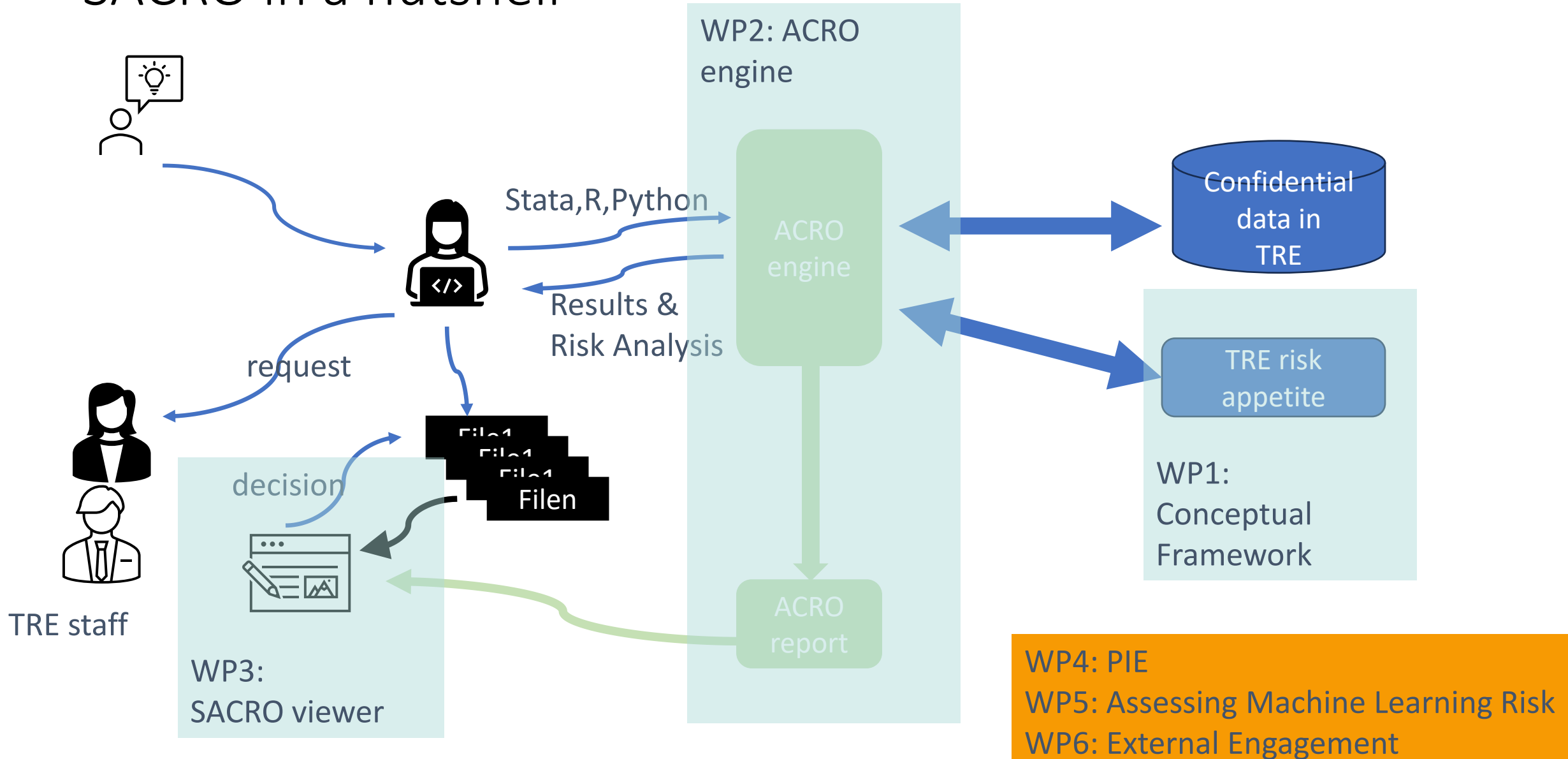
<https://github.com/AI-SDC>

[https://medium.com/sacro-semi-automated-checking-of-research-
outputs](https://medium.com/sacro-semi-automated-checking-of-research-outputs)



Additional Slides

SACRO in a nutshell



SACRO: Semi-automated Checking of Research Outputs

Freeing up TRE staff time

1. Suite of like-for-like replacements for common analysis commands in R/ Python / Stata **that perform disclosure risk assessment in real time**
 - Researchers see SDC feedback immediately => fewer requests for disclosive outputs get made
 - Researchers can add comments/context/ exception requests => fewer rounds of discussion
2. Viewer that shows outputs alongside risk analysis

Establishing Governance Framework for cross-TRE cooperation

3. Consensus statement:
4. Machine & human readable format for specifying risk appetite
 - Vital for federated learning.

Find out more and get involved: sacro.contact@uwe.ac.uk

<https://medium.com/sacro-semi-automated-checking-of-research-outputs>

SACRO Engine

- ✓ Scoping workshops followed by iterated development cycles facilitated TRE involvement in design
- ✓ <https://github.com/AI-SDC/ACRO>
 - ✓ Fully documented,
 - ✓ CI integration
- ✓ Based on industry-standard pandas and statsmodel packages
- ✓ ‘translation wrappers’ for R and Stata
 - ✓ Data passed in memory
 - ✓ “pass through help”
- ✓ Managed to maintain the key ‘drop-in replacement’ philosophy

```
acro.suppress = False
```

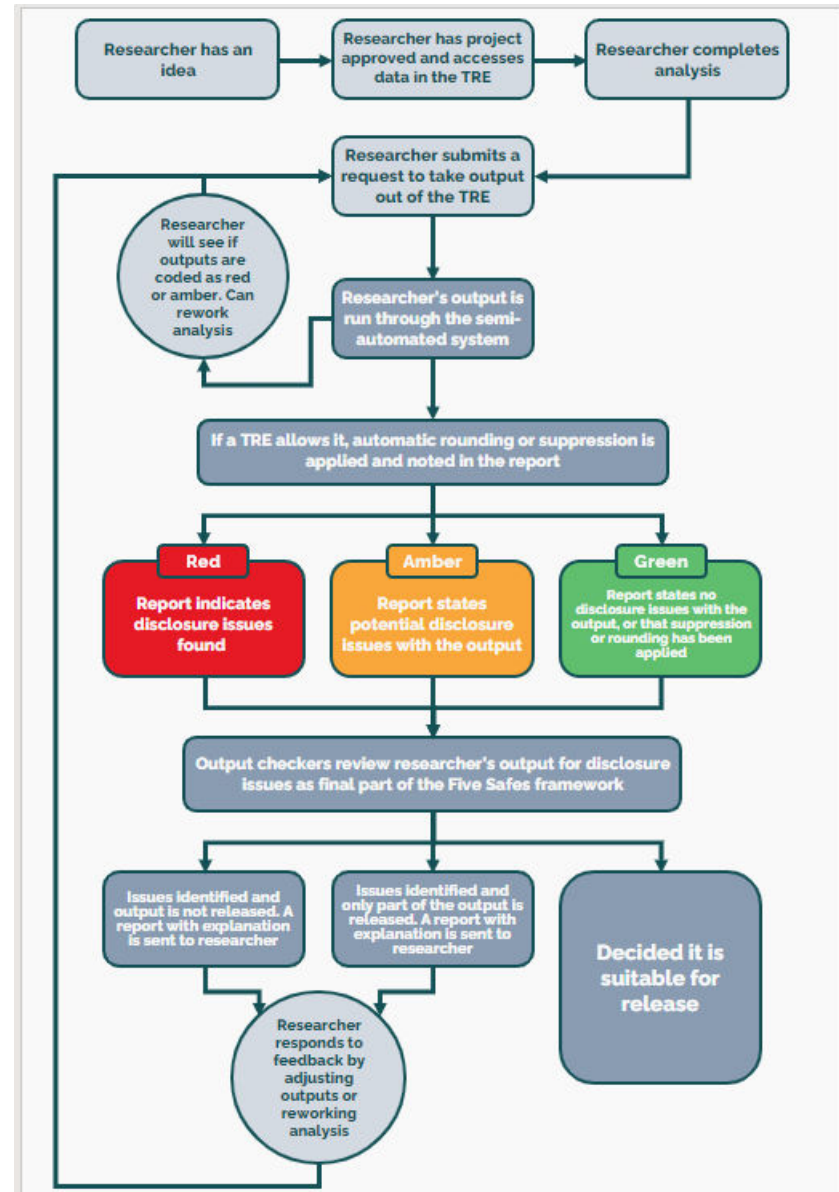
```
safe_table = acro.crosstab(df.recommend, df.parents)  
print(safe_table)
```

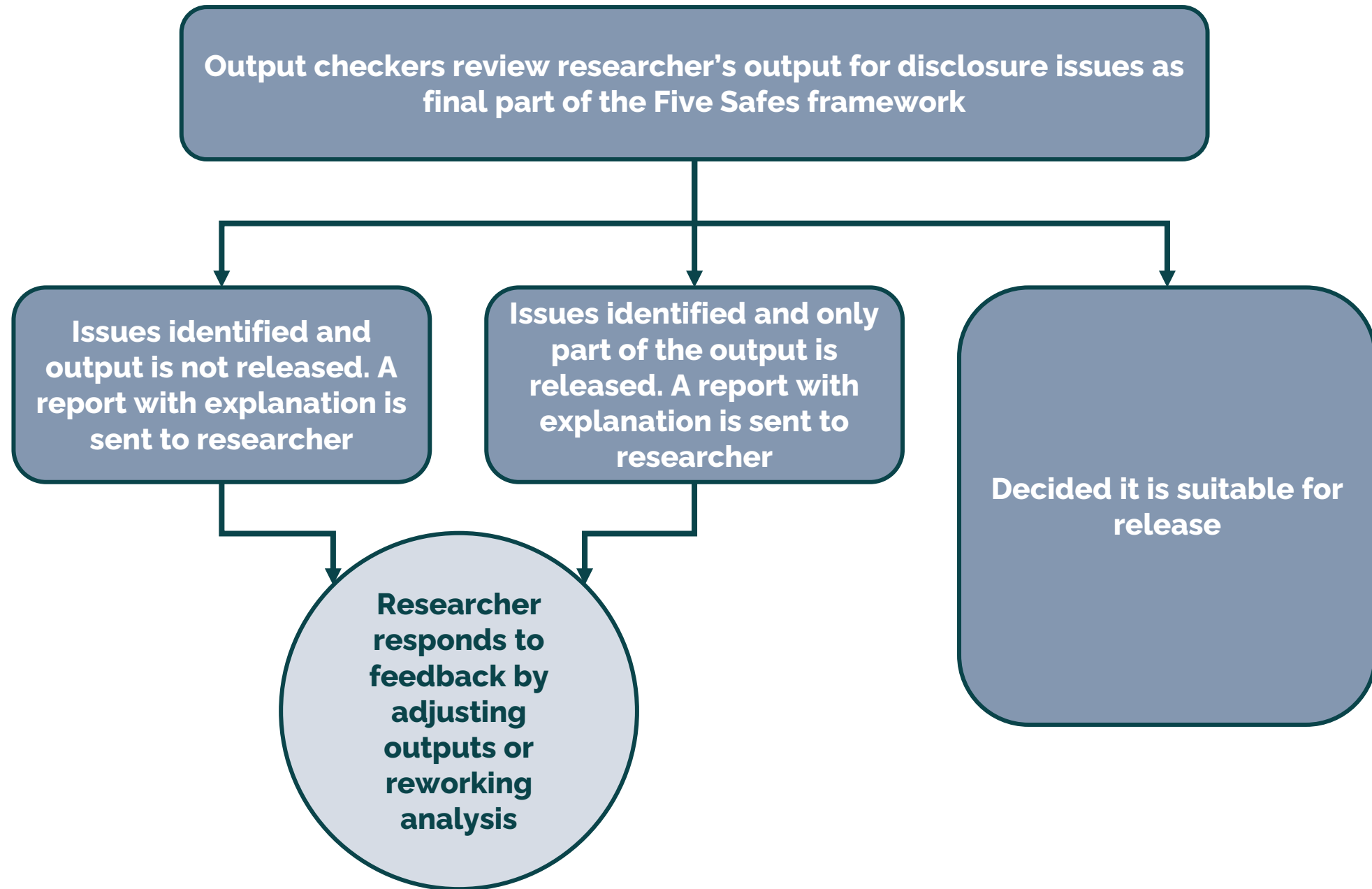
```
INFO:acro:get_summary(): fail; threshold: 4 cells may need suppressing;  
INFO:acro:outcome_df:
```

| parents | great_pret | pretentious | usual |
|------------|------------|-------------|------------|
| not_recom | ok | ok | ok |
| priority | ok | ok | ok |
| recommend | threshold; | threshold; | threshold; |
| spec_prior | ok | ok | ok |
| very_recom | threshold; | ok | ok |

```
INFO:acro:records:add(): output_1
```

| parents | great_pret | pretentious | usual |
|------------|------------|-------------|-------|
| not_recom | 1440 | 1440 | 1440 |
| priority | 858 | 1484 | 1924 |
| recommend | 0 | 0 | 2 |
| spec_prior | 2022 | 1264 | 758 |
| very_recom | 0 | 132 | 196 |





SACRO viewer for output checkers

SACRO Outputs Viewer Release and download

Outputs to review

- output_1 crosstab table
- output_3 pivot_table table
- output_4 ols regression
- output_5 olsr regression
- output_6 probit regression
- output_7 logit regression
- output_8 custom

ACRO risk profile

output_1

Created at: 21 August 2023 at 12:16

Output type: crosstab table

ACRO status: **Fail** (threshold: 4 cells may need suppressing)

Comments:

- Please let me have this data.
- 6 cells were suppressed in this table

Exception request: because I really want it

Review: You cannot approve this output until you add a comment.

Review comments on output_1:

output_1_0.csv Open file

| recommend | great_pret | pretentious | usual |
|------------|------------|-------------|-------|
| not_recom | 1440 | 1440 | 1440 |
| priority | 858 | 1484 | 1924 |
| recommend | 0 | 0 | 2 |
| spec_prior | 2022 | 1264 | 758 |
| very_recom | 0 | 132 | 196 |

Record overall comments, create release package

View TRE risk appetite

Type of output
Recommendation
Comments from user
Exception. request

Comments needed to override recommendation

Outputs to review

Output with problematic cells highlighted