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

October 2023

This consensus statement has been written to encapsulate a set of principles around the use of automated processes when checking that research outputs derived from confidential data do not breach personal privacy.

These principles were derived during the DARE UK funded project SACRO, in consultation with a range of stakeholders, including the public[1]. They apply in both the development and implementation of semi-automated output checking, in Trusted Research Environments (TREs) and other organisations producing outputs from confidential data.

Trusted Research Environments provide a controlled environment where researchers on approved projects can analyse sensitive data such as medical, tax, or census records. Prior to release, research outputs are typically checked by two members of TRE staff who remove anything which could result in the identification of an individual.

To provide decision support, semi-automation can be added to this process to take a 'first look' at each output and advise on its disclosive potential, before outputs are checked by TRE staff. We term this 'semi-automation' as human oversight retains a key role. Such tools will see the process of output checking become quicker and less resource intensive and promote consistency in how TREs manage disclosure control.

Statement

The organisations and individuals who have signed up to this statement support semi-automated checking and the principles of how it should be implemented as outlined below. We would like to see semi-automated output checking implemented more widely in line with these principles.

Principles

The following principles, benefits and considerations have been identified as important when implementing semi-automated output checking:

- Any outputs produced from confidential information held by an organisation, which may be made available beyond that organisation, should undergo a form of output disclosure control checking.

[1] Appendix 1 gives an overview of the process used to establish the principles. Appendices 2 and 3 give more details of the SACRO project, one possible implementation of semi-automated checking.

- There are benefits to semi-automated output checking in making TREs more efficient and less labour intensive, as well as improving data security. Semi-automation allows resources to be concentrated on complex cases, allowing more research to be carried out quickly and securely.

- In addition to supporting output checkers, where there is potential for an automated system to be used to help train TRE staff, this should be utilised.

- It is important to allow TREs to choose their risk appetite and balance within the Five Safes[2]. This applies to all stages of the process including the number of output checkers used alongside the semi-automated system as well as the automatic suppression or rounding used.

- TREs should monitor the error rate in human and machine output checking to inform future plans for automation. This could include monitoring of incidents or how often outputs are stopped at the first output checker.

- Before any semi-automated system is adopted, robust testing of the system is required in TREs across the UK, using a variety of different data.

- Guidance for staff is key to success. Semi-automation will help reduce bottleneck, but output checkers need additional support in the form of a clear guidance document. More generally, training for TRE staff in using the semi-automated system as well as output checking is required.

- Semi-automation need not make use of Artificial Intelligence (AI). However, as such systems develop, future versions may utilise this technology. Consideration should therefore be given to public concerns around artificial intelligence as it relates to disclosure control and how to best communicate this work, as well as how to address issues around biased data sets.

Alongside this statement is an overview of the SACRO project with background, an outline of how public and stakeholder views were collected and details of the findings from these various engagement activities. These findings aim to inform the introduction of semi-automation to output checking. Additionally contained is a summary of the SACRO system and guidance for output checkers.[3][4]

[2] See the beginning of the section titled 'Background' for an overview of the Five Safes.
[3] This work is funded by UK research and Innovation, [Grant Number MC_PC_23006], as part of Phase 1 of the DARE UK (Data and Analytics Research Environments UK) programme, delivered in partnership with Health Data Research UK (HDR UK) and Administrative Data Research UK (ADR UK).
[4] The SACRO project builds on a proof of concept tool 'acro' developed for Eurostat, whose support we gratefully acknowledge.

Call

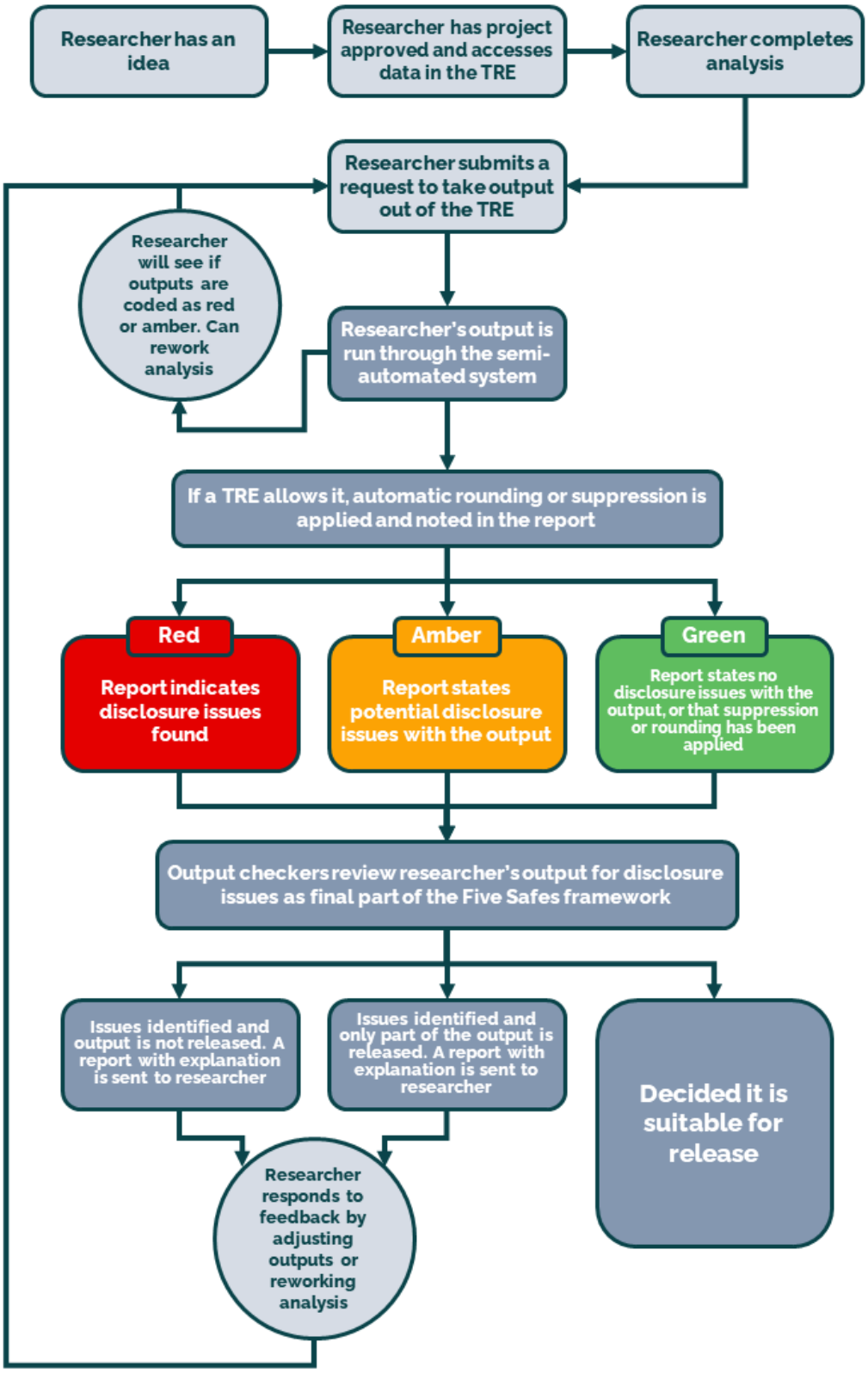
We support semi-automated checking and the principles outlined in this statement that we would like to see implemented more widely. Whilst still in test phase we acknowledge SACRO as an example of semi-automated checking that embodies the principles outlined.

Contact details

If you would like to discuss the contents of this statement, and/or add your support, please contact: Layla Robinson, Chief Partnership & Strategy Officer, at Research Data Scotland (layla.robinson@researchdata.scot).

Overview

The following diagram outlines an overview of how a semi-automated system could work:



Flowchart of the researcher and output checker journey using SACRO. See Appendix 2 and 3 for a full summary of the process.

Registered organisations

The following organisations have agreed to adopt the consensus statement and its principles:

and the Digital Critical Friends group

and the electronic Data Research and Innovation Service (eDRIS)

Contents

Appendix 1: Methodology for Creating Consensus Statement

Background

Trusted Research Environments (TREs)[5] are secure physical or digital environments which can only be accessed by approved researchers. They provide a space in which researchers can analyse confidential data such as health records and report findings.

To keep data secure, the environments are guided by the principles of The Five Safes Framework[6]. These principles are:

Safe People: All researchers accessing the data are trained and accredited.

Safe Projects: Data must be used ethically, for research that delivers clear public benefit.

Safe Settings: The physical and digital settings used to access data are controlled and secured.

Safe Data: Researchers can only access data that has been pseudonymised.

Safe Outputs: All research outputs are checked to ensure individuals cannot be identified.

This consensus statement focuses on the output checking element of this process, and the potential for adding semi-automation to help output checkers make consistent informed judgements and create a quicker and less resource intensive process.

It is worth highlighting that, of the Five Safes, outputs present the lowest risk to data subject confidentiality; nevertheless, it is neither feasible nor desirable to remove human oversight. Outputs can be context-specific, complex and require human oversight[7]. Instead, the aim is to **semi**-automate outputs for which risk has been well documented and guidance established. This is **semi**-automation as it uses and balances both human and automated approaches to output checking.

Currently, outputs are typically checked by two expert staff before release, which is a significant expense for TREs, time-consuming, as well as a potential bottleneck for researchers. Although checking an individual output should be a matter of minutes,

[5] TREs are also sometimes referred to as safe havens or secure data environments (SDEs).
[6] Five safes - Wikipedia, What is the Five Safes framework? | Research Data Scotland.
[7] Alves, K., & Ritchie, F. (2020). Runners, repeaters, strangers and aliens: Operationalising efficient output disclosure control. Statistical Journal of the IAOS, 36(4), 1281-1293. https://doi.org/10.3233/SJI-200661.

the volume of requests means that, in some facilities, it can take ten days for research outputs to be cleared. Additionally, there is a need to create consistency across how different TREs manage disclosure control.

Addressing both these issues, and building on previous work with Eurostat[8], the SACRO project[9] was funded by Data and Analytics Research Environments UK (DARE UK[10]) to investigate and pilot a consistent, efficient, and trustworthy approach to the partial automation of output checking and disclosure control. This is with the ambition to reduce the operating costs of TREs, and the time taken to release research results. The project aimed to:

- Provide guidance to TREs and output checkers to improve consistency.
- Design and implement a semi-automated system for checks on common research outputs. This speeds up output checking so that TREs can choose to reduce the number of human output checkers for low-risk outputs.
- Work with a range of different types of TRE in different sectors (health, social data) and organisations (academia, government, private sector) to ensure wide applicability.
- Work with the public to explore what is needed for public trust and ensure that any automation is acting as "an extra pair of eyes" to support rather than replace TRE staff. By helping them make easy decisions rapidly, their time can be spent on more complex or nuanced cases.

The principles in this document are based on cases in which semi-automation is carried out through a series of rules-based checks. These rules-based checks are human readable and do not generate new knowledge, and so are not considered Artificial Intelligence (AI) within this document[11]. As future semi-automated checking systems could use the data collected to adapt the checks and therefore utilise AI, we reference AI as it arose in public and stakeholder engagement.

UK TREs used 'principles-based' output checking: simple and conservative rules-of-thumb are used to check and approve the bulk of outputs, freeing up resources to provide detailed reviews of 'exceptions' where more detailed scrutiny is justified[12].

---

[8] Green, E., F. Ritchie, and J. Smith (2020). Understanding output checking. Technical report, European Commission (Eurostat - Methodology Directorate).
Green, E., F. Ritchie, and J. Smith (2021, October). Automatic checking of research outputs (ACRO): A tool for dynamic disclosure checks. ESS Statistical Working Papers 2021, 1–27. doi: 10.2785/75954.
[9] Five projects funded to drive more coordinated secure use of sensitive data for research across the UK - DARE UK; SACRO: Semi-Automated Checking of Research Outputs – Medium.
[10] Data and Analytics Research Environments UK (DARE UK) is a programme funded by UK Research and Innovation (UKRI) to design and deliver coordinated and trustworthy national data research infrastructure to support cross-domain research for public good.
[11] See Glossary of Terms.
[12] See also the Glossary of Terms.

SACRO can replace the rules-based part of the process, which is where most resources are currently consumed.

Public and Stakeholder Views

SACRO has conducted a literature review and run a series of workshops with public and industry stakeholders, plus a survey with researchers who use TREs for their work. Additionally, throughout the SACRO project there have been members of the public on the steering group providing input at each stage.

Outlined below is the literature review summary, details of how public and stakeholder views have been collated as well as findings from each.

Literature Review – Summary[13]

The literature review brings together previous public engagement work on output checking, disclosure control, automation and AI, to consider the factors which impact public trust.

The literature consulted showed that the risk of reidentifying individuals in data research was both a recurring concern for people, and an area where their understanding could be greater. This can be seen in the number of questions raised in public engagement work regarding both how effectively de-identification of data can be achieved, and what, in fact, is identifying information. Reidentification is often presumed in the articulation of other fears around data research, too. At the same time, a more liberal attitude to the risk of re-identification has also been recorded, where people are willing to trade off a perceived low risk of re-identification to themselves against the possibility of improving healthcare for others. This altruistic attitude is influenced by participants' trust in the efficacy of the safeguards in place to protect patient data anonymity, which indicates the positive effect that greater awareness has on public trust. This positive effect can be seen in multiple studies that make use of deliberative focus groups, which see concerns reduce and trust increase during the session.

The literature shows concerns around a potential overreliance on automated decision making, as well as an interest in understanding the rationale behind automated decisions. No article was found that dealt with automation in disclosure checking specifically, and the importance of explaining is context-dependent: people feel explanations for AI decisions should be offered in situations where non-AI decisions come with an explanation, apart from in some high-stakes scenarios or scenarios that are more technical than social, such as medical diagnosis, where accuracy is valued over transparency. This raises the question of where output

[13] The full literature review, including references, is available https://medium.com/sacro-semi-automated-checking-of-research-outputs

checking would be seen on this technical/social scale, and the level of explanation required.

The review highlighted public engagement work already undertaken to complement the public engagement which took place specifically for SACRO. However, there are potentially gaps in the public engagement research in this area around future iterations of semi-automation, meaning more work could be undertaken in the future.

Public Engagement

The public engagement activities explored the idea of what matters to the public regarding the disclosure of results. Public trust is essential for TREs to operate, and transparency in changes to processes is part of maintaining this trust. Engagement with members of the public and the creation of a clear document as an output to the project contributes to this transparency. Understanding what builds trust in the processes will inform the future of semi-automation in this field, as well as specifically the delivery of SACRO.

The OpenSAFELY Digital Critical Friends panel at the Bennett Institute and the Scotland Talks Data public panel co-ordinated by Research Data Scotland and Scottish Centre for Administrative Data were consulted on the SACRO project and the contents of this statement. Combined with the literature review findings this has informed a publicly endorsed guideline for statistical disclosure control.

Our public sessions found the following broad themes were important to the public with regards to semi-automation of disclosure control:

- Awareness that data can contain and perpetuate biases.
- Benefits of SACRO, including the idea that automating a process reduces the room for human error.
- The future uses of semi-automated systems, including the training of output checkers.
- The need to manage risk if reducing the number of human output checkers.
- Considered and transparent communication to members of the public.
- The role of artificial intelligence in future semi-automation.

The themes and discussions at the session with the OpenSAFELY Digital Critical Friends were captured in a graphic (included on the following page):

A visual representation of the discussion had by OpenSAFELY Digital Critical Friends in April 2023.

Stakeholder Engagement

Throughout the SACRO project individuals from TREs have been part of the wider steering group and inputted to the stakeholder views as well as informing the direction of SACRO itself.

The SACRO project team held a number of workshops with members of the TRE community to explore relevant issues, such as how output checking is currently undertaken, the use of key programming languages, the bottlenecks, rules and principles-based systems, and what statistical guidance would be useful.

Some broad themes were considered and key points that came through in these discussions are outlined below:

- Output checkers require researchers to include underlying counts in the outputs being checking, the output checker will then ensure that the underlying counts meet the defined threshold.

- Dominance is rarely seen as an issue, it is often a unique problem to business data (where a particular business is the dominant/main employer in a small area, or one business is the main provider of a particular service). Researchers themselves are aware of this issue and how it can impact statistical analysis, so the problem is addressed before analysis is completed.

- TREs generally use a principles, rather than rules-based, approach. It appeared that new analysis could lead to discussions of exemptions, but the most common request for exceptions was due to small numbers or differencing.

- Views on how common exceptions were and the need for guidance varied, but it seemed that exceptions occurred either most weeks or every few weeks and some guidance would be useful. Guidance is key to success. Semi-automation will help reduce bottleneck, but output checkers need additional support in the form of a clear guidance document. More generally, training for TRE staff in using the semi-automated system as well as output checking is required. It should also be clear to researchers how to contact or raise enquiries to TRE staff and where the researchers can find relevant guidance for SDC.

- There was also consideration of what took up output checkers' time and what could be improved. It was discussed that, though an automated solution like SACRO would require some standardisation, it would make things quicker for researchers and reduce errors.

Stakeholder sessions focused on output checkers and TRE operators. Throughout SACRO the researcher experience has also been considered, for example minimal disruption to the usual experience of analysing data using familiar tools with prefixes

(See Appendix 2). Individuals on the SACRO team have experience as researchers required to access data in TREs and have used this experience to inform the build. As TREs evaluated different iterations of the SACRO toolkits, they were asked to 'role-play' as researchers (which many are) and provide feedback from that perspective using SACRO to analyse various open-access datasets. As TREs now move into fuller deployment of the SACRO toolset within their main REE environments we anticipate increasing amounts of feedback on the researcher experience. SACRO have provided TREs (both within and outwith SACRO consortium) sets of questions to understand the researcher experience.

Overall Findings and Outcomes

Themes Which Arose in Public and Stakeholder Engagement

Throughout the consultation with stakeholders and members of the public on the semi-automation of output checking the following themes emerged. These have been taken into account in the development of the SACRO project and its principles and should be considered with future semi-automation of processes.

- It was appreciated that the semi-automation of output checking had a role to play in making TREs more efficient and less labour intensive, thus allowing more research to be carried out. On a similar note, the role a semi-automated system can play in supporting and training output checking staff was also voiced. Public engagement highlighted a belief that training needs to adapt to make sure that when checkers have a green report, they are still applying high concentration levels. The importance of ensuring robust testing of the system in TREs across the UK prior to implementation was raised, too, alongside technical implementation and the need to provide a clear guidance document and training for TRE staff.

- The importance of allowing TREs to choose their own risk appetite, and balance between the Five Safes as they see fit was raised in stakeholder engagement. For instance, if a project has particularly high safeguards regarding the data itself (Safe Data), this reduces the need for the other safes be interpreted as stringently, as all the safes contribute to the same goal of safe research access to data. A semi-automated process should accommodate for this variation.

- When implemented, it was felt that TREs should monitor error rate in human and machine output checking to inform further automation. Looking further ahead, there was also interest in the future of semi-automated output checking and maximising impact. In recognition of this, the SACRO project is looking to the next steps and potential for future development projects.

Overview of the SACRO Process

Using the above findings from public and stakeholder engagement, the SACRO project has developed the following semi-automated process for output checking. More detail is given in Appendix 1.

1. Researchers' analysis will be checked by machine first and graded with a traffic light system:
    - Green – No issues found, or an issue has been dealt with automatically. If the latter, this will be clearly noted in the report.
    - Amber – Potential disclosure issues found.
    - Red – Disclosure issues found.

(NB TREs can choose to use automatic suppression or rounding according to risk appetite. This will be applied before sending reports to the output checker and highlighted.)

2. The output will still then be checked by at least one human output checker. The results from the automated checking are intended to give advice to the checker to support and speed up their decision making and make the system more secure. At the discretion of the TRE, this could reduce the need for two checkers per output to the need for one checker per output.

SACRO does not replace output checking, and piloting and testing are being undertaken with different data sets to ensure its efficacy.

Case Study

A researcher comparing gender, ethnicity and health within a geographical area might generate a table with those factors as part of their analysis. This table would then be submitted to the SACRO system which runs checks on all the files including this table and produces a report. In this instance, there is only one individual who is defined as Asian and Male who has suffered a heart attack, so the SACRO system would flag this as Red, as it could be possible to identify that individual.

A human output checker would then have to review the file and report back to the researcher the reason why it was rejected. The researcher then re-works their analysis, perhaps re-grouping the individuals to create a new 'Other' group or suppressing the identifiable information.

Alternatively, if the TRE has enabled automatic suppression, this will be applied before a 'Green after automatic suppression applied' report is sent to output checker, who will review before releasing.

Guidance

In addition to implementing a semi-automated system in TREs, it was identified as important to produce a guidance document to support consistency for output checkers and outline how to work with a semi-automated system. This guidance will be available via the SACRO Medium page.[14]

[14] SACRO Medium webpage SACRO: Semi-Automated Checking of Research Outputs – Medium

Appendix 2: The SACRO Project

This project builds on existing work in this area and brings together the latest advances to provide framework and resources (including practical guidelines) to support disclosure control in different technical and procedural environments. To achieve efficiency and transparency, we have created resources that help researchers make fewer disclosive requests and TREs spot these faster. The design framework and tools:

- Support researchers to use the major analytical languages (R, Python and Stata), with minimal changes, by exploiting the 'wrapper' approach we have successfully trialled elsewhere.
- Support TREs with different operating models and output checking workflows, through a process of co-design to maximise useability.
- Automate checking of most common statistics, using best-practice principles-based modelling.
- Build on the GRAIMATTER[15] project to create practical guidelines and prototype tools for triaging AI models for clearance.

SACRO Process

Researchers carry out their analysis using familiar tools in R/stata/python etc, the only change being slightly different prefixes so that lightweight 'wrapper' translation functions are called.

For example, to make a logit model in R, they pass the relevant formula and data to acro glm instead of glm()

Back-end code automatically inserts appropriate automated disclosure check such as thresholds, dominance, etc., before passing query to standard libraries.

For transparency, the researchers are given the statistical disclosure control (SDC) output alongside their analytic results, so that they can avoid asking for disclosive outputs, to reinforce their training about SDC.

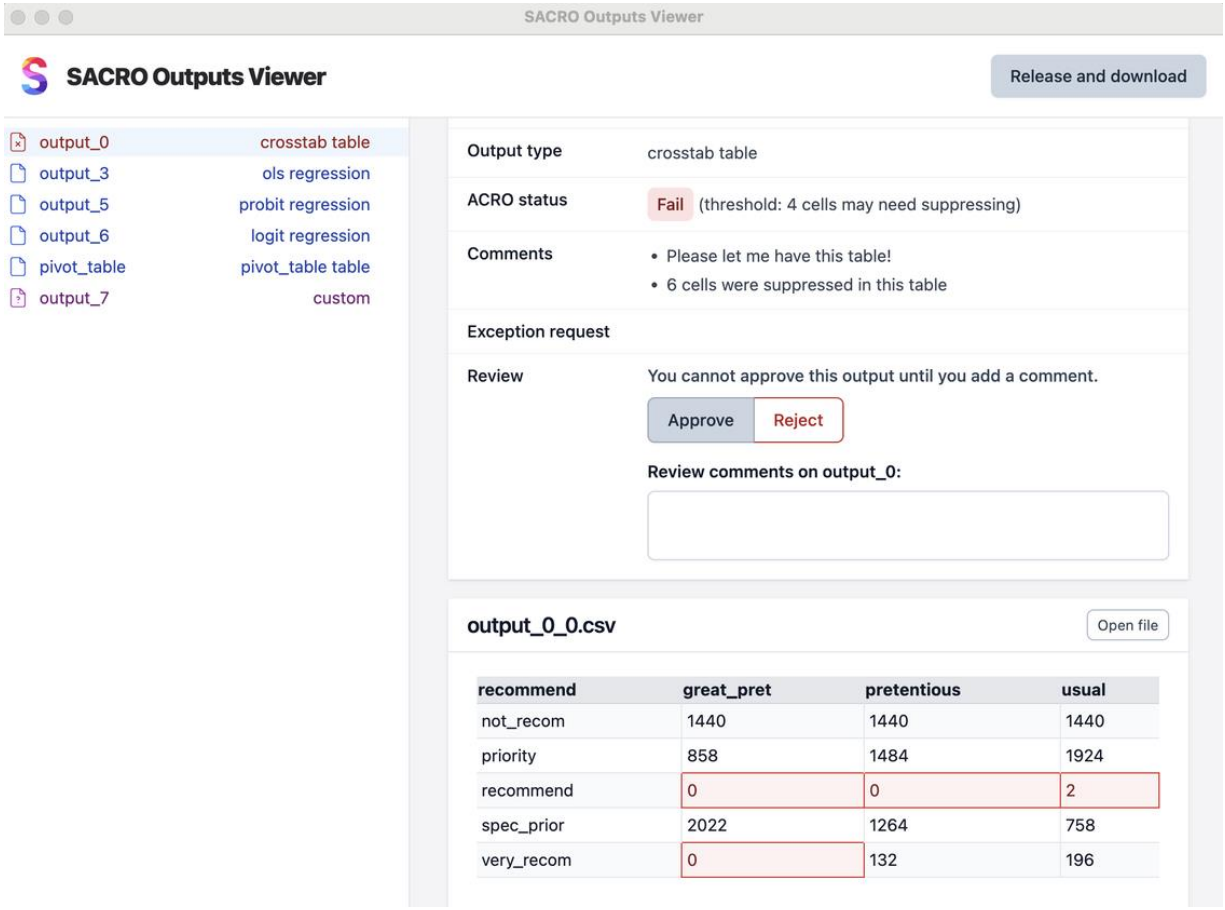
When ready, they call a 'finalise() function, which puts their requested outputs in the checking process.

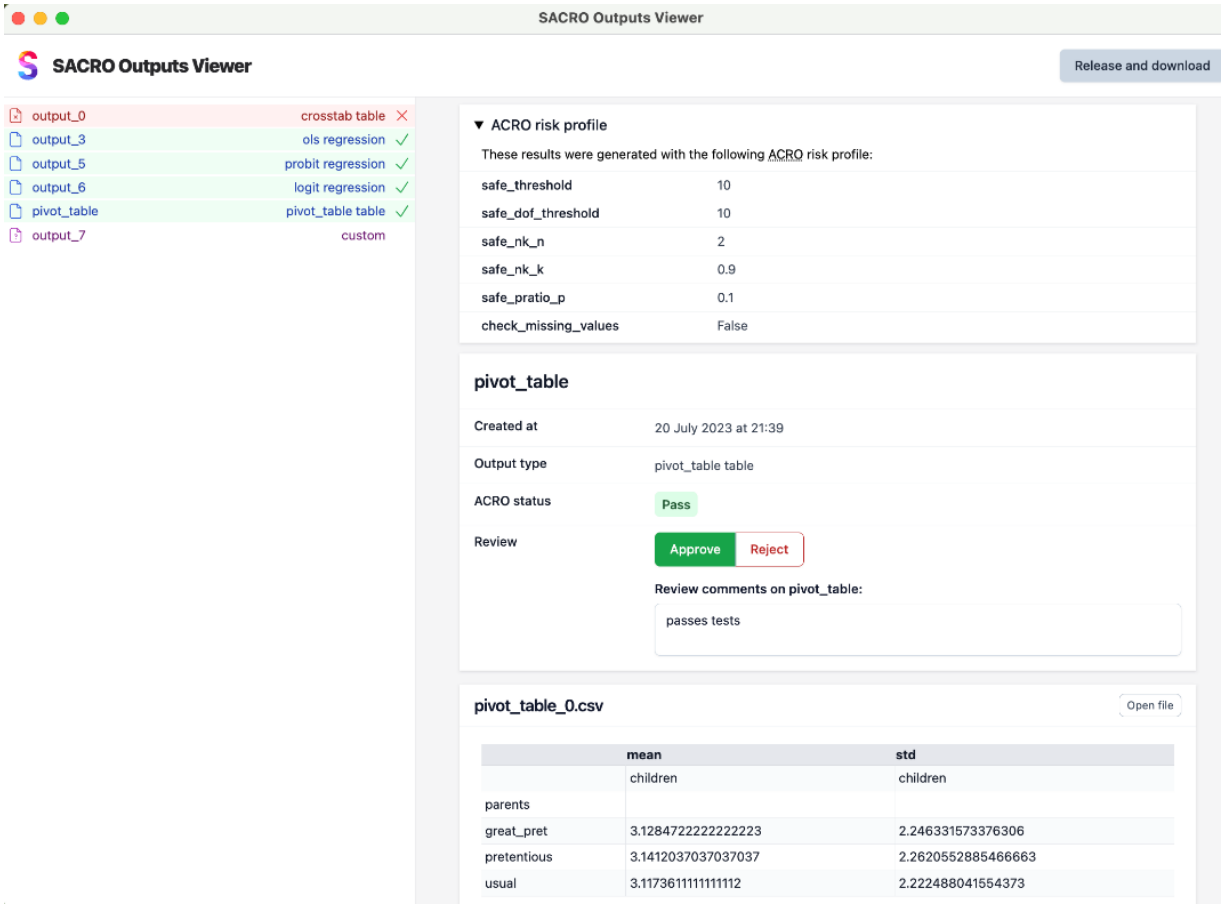
[15] GRAIMATTER: Guidelines and Resources for Artificial Intelligence Model Access from Trusted Research Environments - DARE UK

A different version of the user interface lets the TRE staff view outputs, risk analysis, and the TRE risk appetite, and so make faster better decisions.

Application Programme Interfaces (APIs) will be designed to maximise the possibility that these stand-alone tools can be integrated into the workflows at different TREs.

On the following page is an example of the reports an output checker will see after submitting through the SACRO system. The left-hand column shows a list of files requested. In the top image, colouring of file names suggests which files require special attention. In the lower image, background colour-coding and tick/cross symbols show decisions made by output checker. The top image shows checker viewing table that fails disclosure tests, with problematic cells highlighted in red. The lower image shows an acceptable table. In that same image, the top right-hand panel shows the option to view TREs 'risk appetite' expanded.


SACRO Outputs Viewer
SACRO Outputs Viewer
Release and download
output_0 crosstab table
output_3 ols regression
output_5 probit regression
output_6 logit regression
pivot_table pivot_table table
output_7 custom
Output type
crosstab table
ACRO status
Fail
(threshold: 4 cells may need suppressing)
Comments
• Please let me have this table!
• 6 cells were suppressed in this table
Exception request
Review
You cannot approve this output until you add a comment.
Approve
Reject
Review comments on output_0:
output_0_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


SACRO Outputs Viewer
SACRO Outputs Viewer
Release and download
output_0 crosstab table
output_3 ols regression
output_5 probit regression
output_6 logit regression
pivot_table pivot_table table
output_7 custom
ACRO risk profile
These results were generated with the following ACRO risk profile:
safe_threshold 10
safe_dof_threshold 10
safe_nk_n 2
safe_nk_k 0.9
safe_pratio_p 0.1
check_missing_values False
pivot_table
Created at
20 July 2023 at 21:39
Output type
pivot_table table
ACRO status
Pass
Review
Approve
Reject
Review comments on pivot_table:
passes tests
pivot_table_0.csv
Open file
mean
std
children
children
parents
great_pret 3.1284722222222223 2.246331573376306
pretentious 3.1412037037037037 2.2620552885466663
usual 3.1173611111111112 2.222488041554373

Appendix 3: Case Study

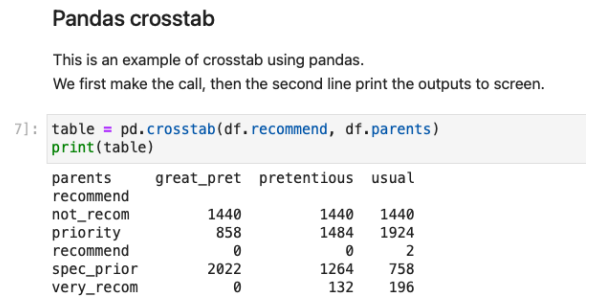
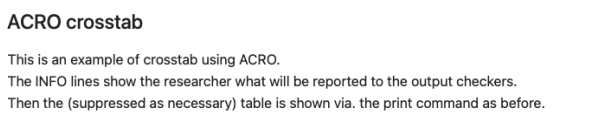
This is an example of how the SACRO system would work on a data set about admissions to a nursery school. This example has been run using an open-source data set[16]. Included are screenshots of what the researcher sees, both with and without using the ACRO command, as well as the output checker's view.

Researcher using the ACRO command

Once a researcher has access to the data, they will analyse it using commands they are used to using. For example, shown below is a request for a table of types of parents admitting their child to the nursery school.

If using SACRO, they will add an ACRO command to the start of their code. When first running ACRO, they are shown the risk threshold.

Shown below are two examples of what the researcher would see without and with using ACRO for a request that fails a threshold check.

Without ACRO	With ACRO
In original (PANDAS) researchers would have to manually scan for numbers under the threshold.	With ACRO researchers get the results in a table which explicitly flags the results as being under threshold. They receive an additional table (highlighted in pink below) as well as the original table.
	

[16] https://www.openml.org/search?type=data&sort=runs&id=26&status=active

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

If a request fails more checks, then this is shown in the table.

With ACRO

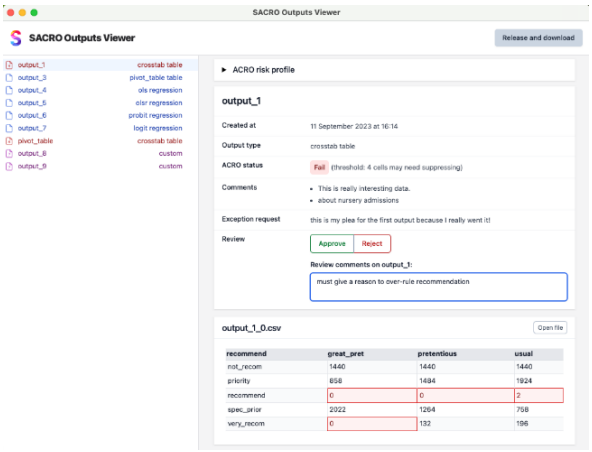
```
INFO:acro:outcome_df:
-------------------------------------------------------------------------------------------|
parents    |great_pret          |pretentious         |usual                                 |
recommend  |                    |                    |                                      |
-------------------------------------------------------------------------------------------|
not_recom  |                 ok |                 ok |                                  ok|
priority   |                 ok |                 ok |                                  ok|
recommend  | p-ratio; nk-rule;  | p-ratio; nk-rule;  | threshold; p-ratio; nk-rule;  |
spec_prior |                 ok |                 ok |                                  ok|
very_recom | p-ratio; nk-rule;  |                 ok |                                  ok|
-------------------------------------------------------------------------------------------|
```

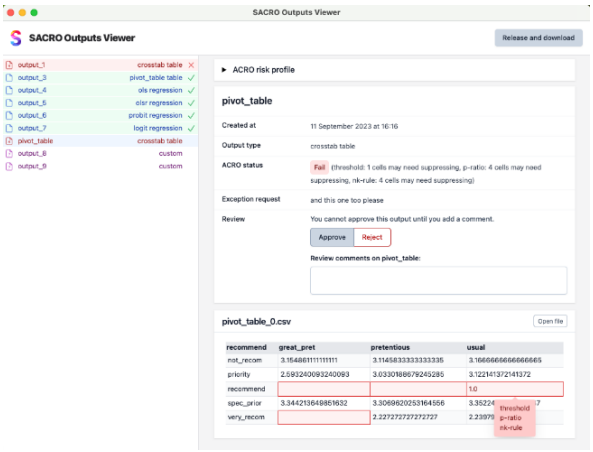
Here, three checks have been failed: threshold, p-ratio and k-ratio.

Researcher submitting results for checking

Once the researcher finalises their outputs, they are prompted to add comments to the outputs, which are flagged as failing checks. They can either write a justification for the output being released, or they can comment that they do not need that particular output.

Output checker reviewing submissions





Once the researcher submits their outputs, the output checkers can view results in the SACRO Outputs Viewer. Researchers can also submit unsupported data, which cannot be automatically checked by the SACRO tool, and then the output checker can manually check it within the same system.

Output checkers review and approve or reject each output and can add comments. Output checkers then finalise their review and add any final comments for the submission as a whole.

A zip file of only the approved outputs is sent to the researcher to download outside the TRE. They also receive a text file with the comments from the output checker. Zip files will contain all the outputs in standard CSV file format.

```
# Review summary
All outputs cross-checked for differencing risk.

## output_1
ACRO status: fail
Reviewer status: rejected

Insufficient reason given to over-rule recommendation in this case.

## output_3
ACRO status: pass
Reviewer status: rejected

This seems like unnecessary precision and  might disclose the cohort size?

## output_4
ACRO status: pass
Reviewer status: approved

## output_5
ACRO status: pass
Reviewer status: approved

## output_6
ACRO status: pass
Reviewer status: approved

## output_7
ACRO status: pass
Reviewer status: approved

## pivot_table
ACRO status: fail
Reviewer status: rejected

## output_8
ACRO status: review
Reviewer status: approved

This image file does not contain disclosve data.

## output_9
ACRO status: review
Reviewer status: approved

This is the code - I have checked that it does not
contain record-level data.
```

Appendix 4: Glossary of Terms[17]

Term	Definition/More information
Artificial Intelligence (AI)	Artificial intelligence is the intelligence of machines or software, as opposed to the intelligence of human beings or animals. This is defined here as a process which generates new knowledge based on the data collected.
Application programming interface (API)	An API is a piece of software which allows different computer applications to communicate easily and securely.
Human Readable rules	Rules which a person, as opposed to a computer, can read and understand.
Rule Based Checks	These are checks which are carried out by a set of human readable rules. The rules do not adapt or change unless edited by a human.
Principles-based output checking	An approach to output checking which uses rule-based checks to arrive at a preliminary decision to release or not but reserves the right to review this decision if the case to do so can be made by the researcher or facility manager. By placing the subjective assessment of risk at the forefront of the decision to release an output, a high level of expertise is required from the decision maker.

[17] GRAIMATTER: Guidelines and Resources for Artificial Intelligence Model Access from Trusted Research Environments - DARE UK

Semi-automated	A fully automated process does not involve any human oversight or input. SACRO is a **semi**-automated process, as human oversight/judgement is still required on each output after the automated element of the process has occurred.
Statistical Disclosure Control (SDC)	Researchers will agree that any outputs from their analysis will not be able to identify an individual – therefore researchers should not attempt to publish any output that immediately identifies individuals (e.g. as an outlier). This is known as disclosure. One of the final steps for a researcher using any personal or individual level data is to undertake statistical disclosure control (SDC) on research outputs - this means using statistical techniques to make the outputs less identifiable.[18]
Trusted Research Environment (TRE) or Safe Havens or Secure Data Environments (SDEs)	The term Trusted Research Environment (TRE) can be used interchangeably with Safe Haven and Secure Data Environment (SDE). We have generally used TRE in this paper for consistency. TREs are highly secure computing environments that provide (remote) access to data for approved research.
Output checking	This is the process by which TRE staff check files which researchers would like to export from the TRE to ensure that they do not contain any potentially identifiable data. Statistical disclosure control Steps taken with data to eliminate (or reduce) the risk of

[18] ADR Scotland (2021), Researcher handbook: Data linkage and administrative data research in Scotland

	disclosing information about a person from the data.
Output checkers	These are trained staffed based within TREs who check files researchers wish to export from the TRE. They decide if files can be released or not.

This document was developed as an output of the project SACRO (Semi-Automatic checking of Research Outputs). This work is funded by UK research and Innovation [Grant Number MC_PC_23006] as part of Phase 1 of the DARE UK (Data and Analytics Research Environments UK) programme, delivered in partnership with Health Data Research UK (HDR UK) and Administrative Data Research UK (ADR UK).