

Implementing the Five Safes Framework at the Bundesbank RDSC through Automation

Pol van Rijn, Sebastian Seltsmann,
Christian Hirsch, Stefan Bender*

* This project would not be possible without the expertise and experience of all colleagues in the RDSC.

The expressed views do not necessarily reflect the opinion of the Bundesbank.

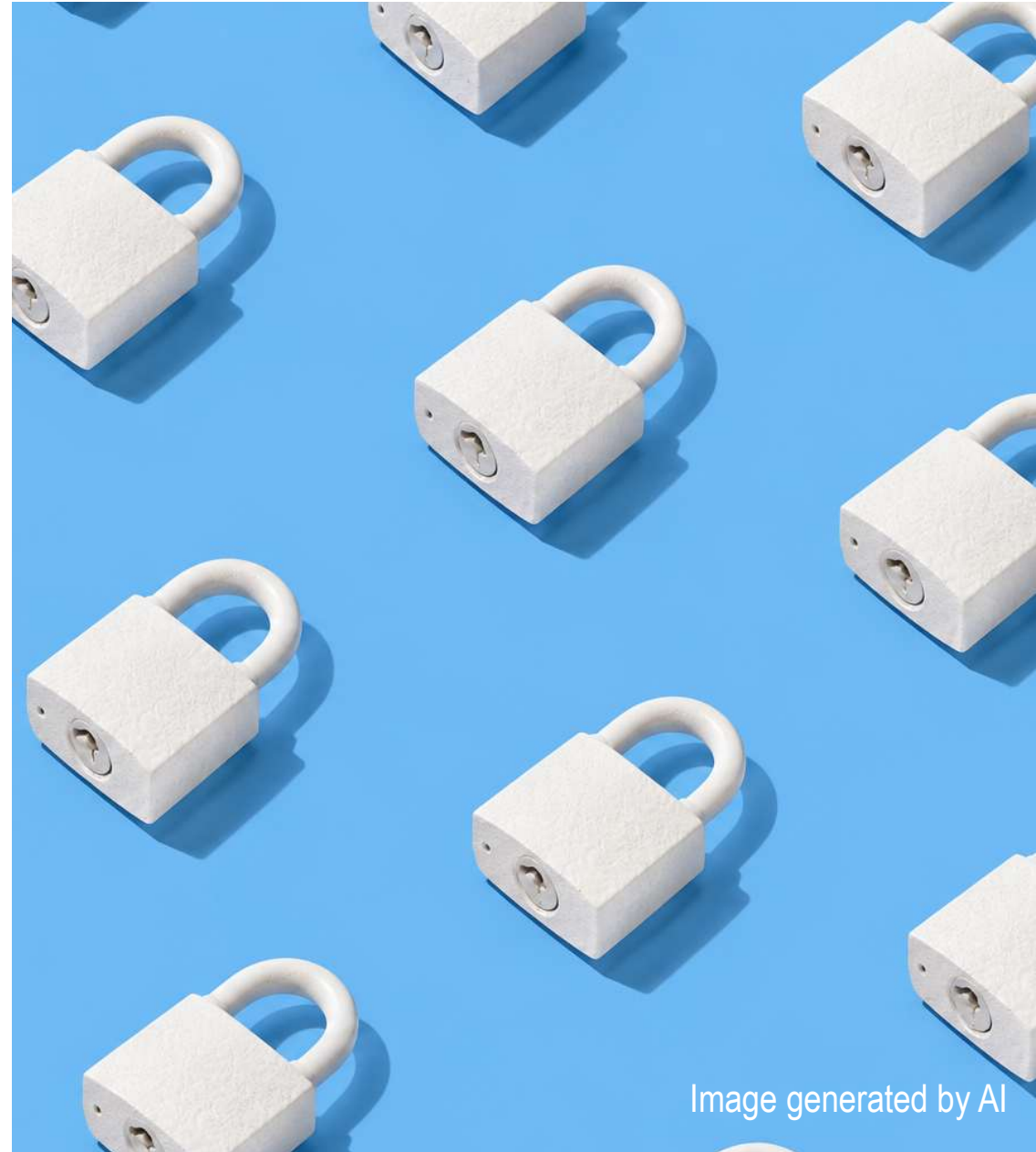


Image generated by AI

Secure access as a workflow challenge

- RDSC provides controlled access to confidential microdata
- Secure access requires **legal, technical, and organizational safeguards**
- They implement the **Five Safes framework**:
 -  **Safe Projects:** Assess per application: purpose, legal basis, public interest, requested data, and access conditions
 -  **Safe People:** Researchers are identified, qualified, trained, and bound by confidentiality
 -  **Safe Settings:** Access is restricted to controlled RDSC environments and defined technical access arrangements
 -  **Safe Data:** Access is limited to approved data products
 -  **Safe Outputs:** Results undergo disclosure-risk review before leaving secure env.
- **However**, mainly implemented through **manual, fragmented workflows**
- ▶ **Consequences:** Duplication, high coordination effort, avoidable delays, and limited transparency → **Beneficial for all parties involved**

Four tools to operationalize the Five Safes

- ▶ Our aim: Automate routine tasks while retaining human accountability
- Four connected tools in different parts of the project lifecycle:
Addressing specific operational weakness in the current implementation

Tool	Current challenge	Automation approach
Application form 	PDF with limited validation and follow-ups required	Browser-based form producing self-contained application file
Project workflows   	Rules, responsibilities, statuses scattered across systems	Web application as central workflow system
Data reports  	Isolated documentation per data release	Connected, browsable HTML data universe (online or offline)
Reproduction   	Require substantial manual coordination	Researcher-triggered replication run using TRACE

Application form

- Static HTML form (Node.js) producing self-contained ZIP including all documents
- ✓ Works in any browser (no software required)
- ✓ No server/DB required (no security risk)
- ✓ Advanced validation (including follow-ups)
- ✓ One application bundle (required documents)
- ✓ Completely open-source (public example)

 [rdsc_form.html](#)  [public_example.html](#)

BdP-IFC Workshop: Five Safes through Automation



Research Data and Service Centre 

Research Project Application

[Clear form](#)

Continue an existing application

Drag an application ZIP here, or choose a file.

 Choose ZIP

Data access

Data access reason

Select...



Project

Project information

Title

Research question and hypotheses

Word count: 0

Briefly explain and motivate your research question, ending by clearly stating your theoretical hypotheses (approx. 150–300 words).

Design principles for new application form

- **Fail fast:** Avoid wasting time on projects that cannot be approved
- **Increased transparency:** List all datasets which can be used and how they can be linked
- **Single submission:** One application per project, all required documents sent with the initial submission
- **Structured input:** Constrained options instead of free text answers
- **Think of the general case first:** Start with a public example, used as boilerplate

Required documents

Attach the

- Curriculum
- Anford
- Isx

Print, sign

- Me

Continue an existing application

Drag an application ZIP here, or choose a file.

Data access

Data may only be provided when the research benefits the public or the scientific community. Please tell us the purpose first.

What is the reason for accessing the data?

Select...

Project

Project title

Goal of the project

Word count: 0

Only indirect
(matching on rates are usual)

List projects

https://poseidon-test.inet.bundesbank.de/de/projects?datasets=B-Daten&order_by=datasets&order_direction=desc

 Poseidon

 Projekte

 Data

 People

 Procedures

 Tasks

 Mitteilungen

 Suche

 Projects

Add new

10

Zeilen anzeigen

Suche:

Search across all fields

Project	Project Title	Lead researcher	Datasets	Status	Supervisors	Last activity
Search Projec	Search Project Title	Search Lead r	B-Daten	Search Stz	Search Supervis	
 2025_1011	Sufficient Statistics for Bank Capital Regulation	 Maxi Miliano	 GuV  B-Daten  BISTA DOM  JANIS	Running	 admin	1 Woche, 2 Tage her
 2025_1005	Automatisierte Betrugserkennung im Zahlungsverkehr	 Emily Carter	 GuV  B-Daten	Running	 Tobi	1 Woche, 2 Tage her
 2025_1001	Auswirkungen von KI-gestützter Prognose auf das Risikomanagement während der Corona-Pandemie	 Emily Carter	 GuV  B-Daten	Running	 Tobi	1 Woche, 2 Tage her

1 bis 3 von 3 Einträgen (gefiltert von 12 Einträgen)

«

<

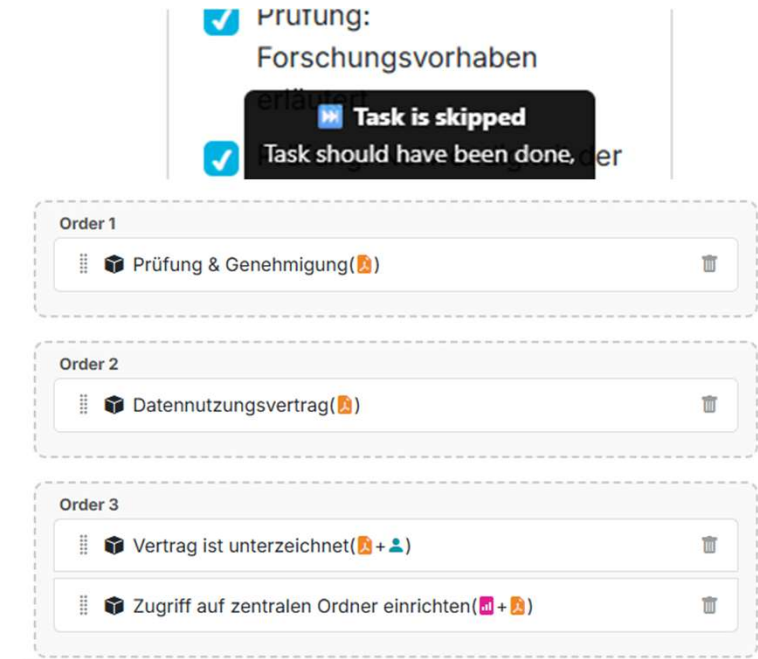
1

>

»

Poseidon: Our project workflow tool

- As rigid as possible, as flexible as needed:
 - Standardize workflows: clear what's done next
 - Allow deviations from rules when necessary, but keep them explicit and documented
- **Reusability:** Avoid repetition, reuse steps that have to be done for multiple datasets
- **Extendability:** If rules change, projects update automatically
- **Automate as much as possible:** Replace manual triggers with automatic triggers where feasible



From isolated reports to a data universe

WIP

- **Data reports** are standardized, technical documents fully describing dataset and variables
- Our tool extracts and structures **dataset schemas** and **descriptions** in a machine-readable form enabling
 - ✓ **Compare releases** Identify exactly what changed (even when minimally updated)
 - ✓ **Connect datasets:** Categorize variables by meaning (`cty_lender` → place > city)
 - ✓ **Generate dummy data:** Declared schema makes the expected data structure explicit
 - ✓ **Automated support:** Machine-readable descriptions enable tools and AI agents to reason about available data
- ▶ **Structure first, views second:** Once the data structure is correct, multiple views become straightforward
- ▶ **Show only what is needed:** Navigate from overview to detail rather than presenting everything at once

TRACE: automatic reproducibility



- **TRACE** allows to automatically reproduce researchers' on-site work on microdata
- We developed a **MVP** within the existing RDSC infrastructure:
 - Researchers push code to a dedicated replication server
 - The server retrieves approved data using the researcher's existing credentials
 - Results can be verified without moving confidential data outside the protected setting
- The Bundesbank's transition to a **cloud-based data platform** requires adaptations for operational deployment
 - ▶ **Researcher-triggered, staff-light:** Reproduction should run without additional manual coordination by RDSC staff
 - ▶ **Respect existing controls:** Use existing researcher credentials
 - ▶ **Keep code and data separate:** Also a first step towards remote execution

Summary and outlook

- We presented four tools under development at the Bundesbank RDSC:
 - **Application form** addresses incomplete and invalid project submissions
 - **Poseidon** connects fragmented information, requirements, and evidence into a project workflow tool
 - Our **data report tool** turns disconnected PDFs into structured, connected documentation
 - **TRACE** automates reproduction of results in the secure environment
- Together, not only staff but also researchers benefit from this
- **Next challenge:** Automated support for output disclosure checking under **Safe Outputs**