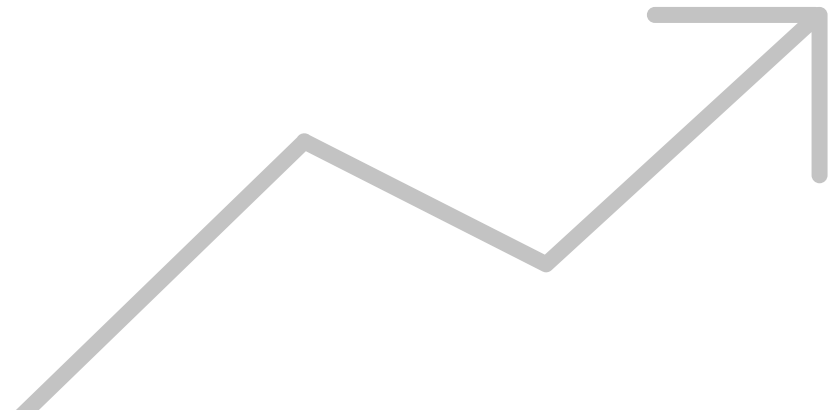


Remote Access for Scientific Use Files – a New Pathway for German Official Statistics Microdata Access

UNECE - Expert Meeting on Statistical Data Confidentiality

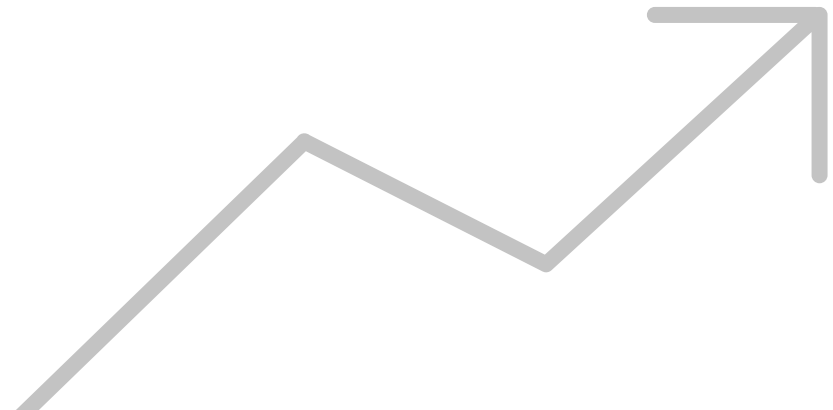
26-28 September 2023, Wiesbaden

Hanna Brenzel, Katharina Cramer, Volker Güttgemanns, Marcel Mathes, Hariolf Merkle

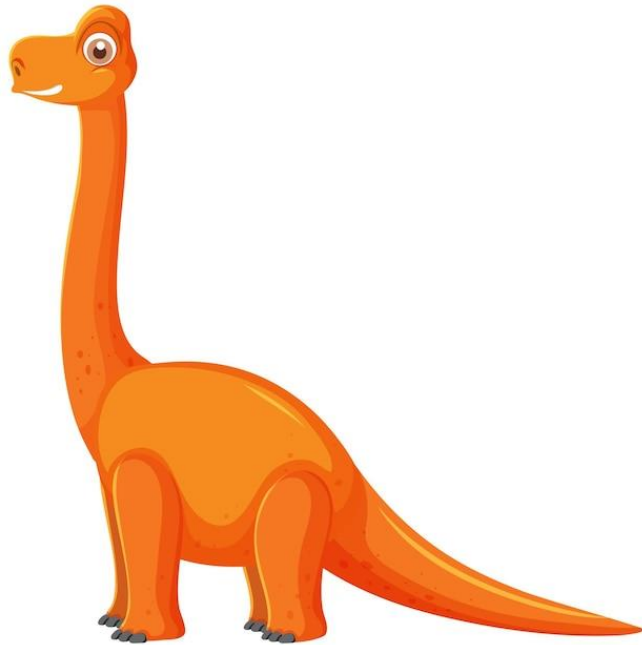


Agenda

- (1) Motivation
- (2) Status Quo
- (3) The Remote Access System
- (4) Outline



Motivation



Remote access for SUF...

- offers convenient data access for scientists from their own scientific institution and thus enables up-to-date and efficient data analysis
- offers the scientific community the opportunity to save travel and waiting time
- contributes sustainably to the further development of the RDC and its range of services
- drives the digitization of procedures and processes
- promotes the awareness of confidentiality requirements by the scientific community and favors faster statistical confidentiality checks, provision and publication of results

Status Quo

	Anonymisation			
	On-site use		Off-site use	
Way of access	Remote execution	Safe centres	Off-site Scientific Use Files	Public Use Files/ Campus Files
Degree of data anonymisation	Formally anonymous	Formally anonymous	De facto anonymous	Absolutely anonymous
Entitled to use	Research institution	Research institution	Research institution	All
Data storage during use	Statistical offices	Statistical offices	Research institution	Any
Location of users during use	Any	Location of the RDC	Research institution	Any
	Analysis potential			

Where to go...

	Anonymisation				
	On-site use		Under construction	Off-site use	
Way of access	Remote execution	Safe centres	Remote access	Off-site Scientific Use Files	Public Use Files/ Campus Files
Degree of data anonymisation	Formally anonymous	Formally anonymous	De facto anonymous	De facto anonymous	Absolutely anonymous
Entitled to use	Research institution	Research institution	Research institution	Research institution	All
Data storage during use	Statistical offices	Statistical offices	Statistical offices	Research institution	Any
Location of users during use	Any	Location of the RDC	Research institution	Research institution	Any
Analysis potential					

Remote Access System

Technical and organizational measures

Technical measures

- » Transport encryption
- » Two-factor authentication IP whitelisting
- » Server-side monitoring/logging/backup
- » Operation in high security environment
- » Limited scope of use and user-specific working accounts
- » BSI-compliant hardware and software administration

Organizational/contractual measures

- » Clause on up-to-date software on the hardware of the facilities
- » Contractual ban on re-identification
- » Exclusion of the institution in case of misconduct
- » Clause on location-based access to data
- » Access only for eligible institutions according to Section 16 (6) Federal Statistical Act
- » Fines for breach of contract

Anonymisation measures

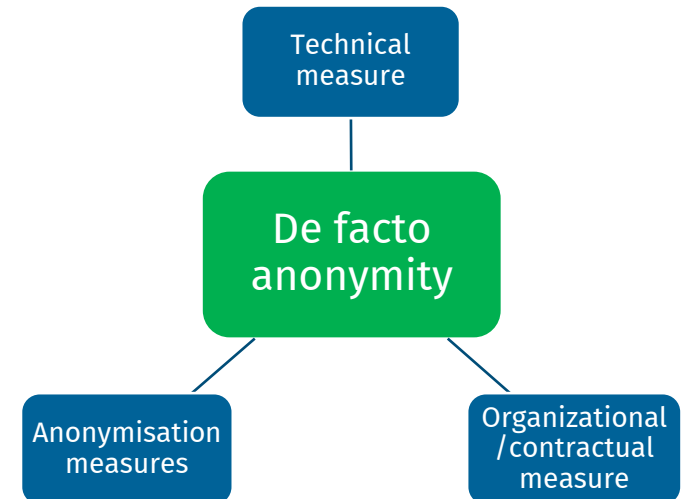
- » Anonymisation concept and special anonymisation of vulnerable units

Remote Access System

Data material (I)

“de facto anonymity”

- » “requires unreasonable effort in terms of time, cost and manpower” (Section 16 (6) of the Federal Statistics Act)
- » de facto anonymity is thus not a fixed quantity, but can be mapped along a continuum
- » the higher the technical and contractual measures, the fewer anonymization measures need to be taken and the higher the analysis potential of the data

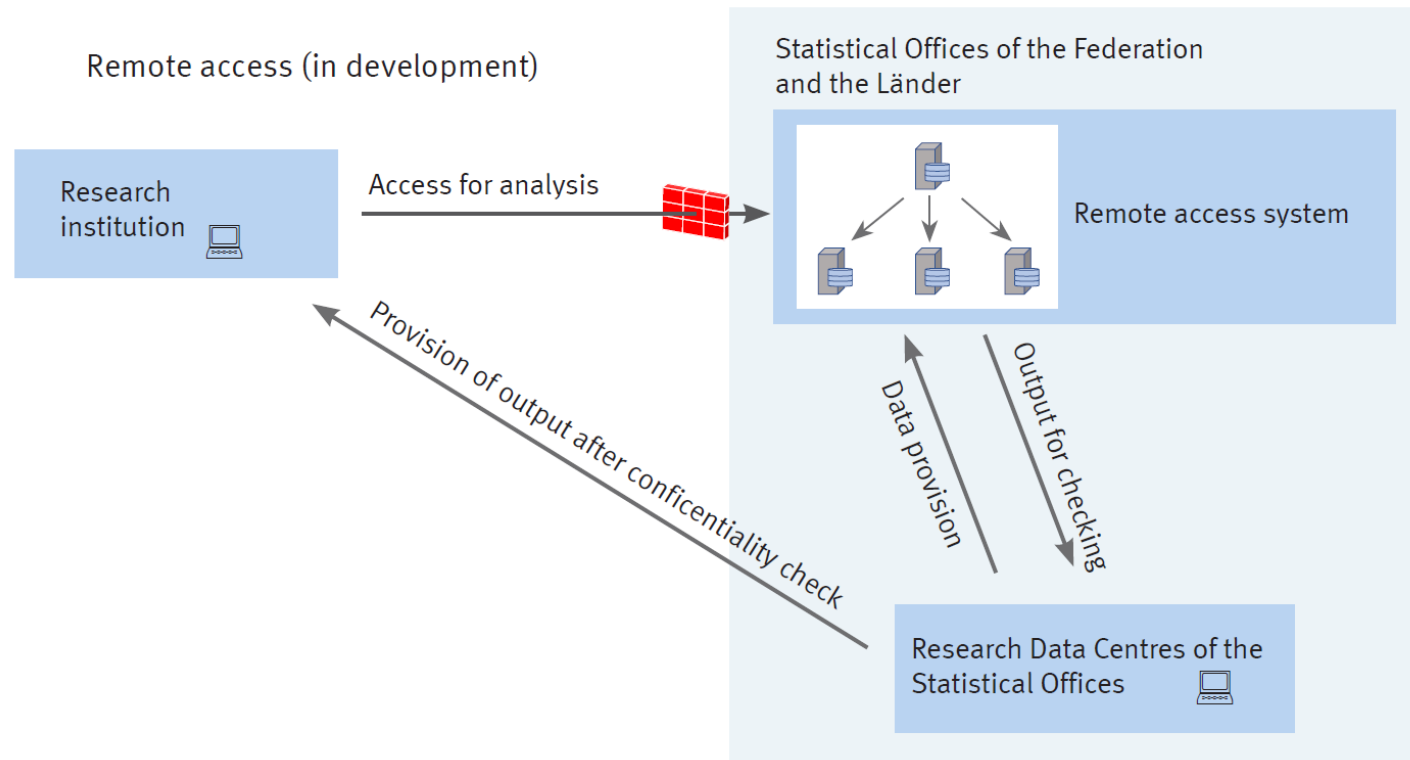


Remote Access System

Data material (II)

- » Significant difference to the current dissemination path via off-site SUF
- » With the new remote SUF, de facto anonymity can be achieved by less modification of the data. This is justified by the high level of technical measures and the associated possibility to control the data access
- » In contrast to off-site SUF, the data is not passed on. It is solely possible to *view* the data via a virtual desktop
- » Unauthorized data linkage is impossible and the RDC has a high level of use control compared to off-site SUFs

Remote Access at the RDC



Source: German version published in WISTA | 6 | 2022, p. 54 et seq.

Outline

- » Phase 2: Evaluation of the system
 - » multi-stage application and function tests for various user groups
 - » penetration, accessibility, load and performance tests
 - » feasibility and usability of SecureBootSticks
- » System dimensioning for the desired number of usage accesses
- » Expansion of the offered data materials

Contact

Research Data Centre
of the Federal Statistical Office

Phone: +49 611 / 75-2420

E-Mail: forschungsdatenzentrum@destatis.de

www.forschungsdatenzentrum.de

