Remote Access in verschiedenen Forschungsdatenzentren – Stand der Technik

Sebastian Ellert, Berlin

Dieser Beitrag untersucht die Herausforderungen und Lösungsansätze der Datenfernverarbeitung in wissenschaftlichen Forschungsinstitutionen. Hierzu werden Begriffe wie On-Site- und Off-Site-Nutzung, Remote Access, Remote Execution und Remote Desktop definiert. Bei der Beschreibung technischer Umsetzungsmöglichkeiten wird besonderes Augenmerk auf Sicherheit und Flexibilität gelegt, wobei Safe-Rooms und Gastwissenschafts-Arbeitsplätze einbezogen werden. Internationale Beispiele wie MONA und Secure Lab zeigen Gestaltungsmöglichkeiten innerhalb der verschiedenen Systeme. Abschließend werden für Institutionen, die bisher keinen Remote Access anbieten, Empfehlungen zur Verbesserung von Effizienz und Sicherheit ausgesprochen.

1. Motivation

Daten sind der Schlüssel zu neuen Erkenntnissen. Insbesondere Verwaltungsdaten können wertvolle Einblicke in gesellschaftliche, wirtschaftliche und politische Zusammenhänge geben. Um den größtmöglichen Nutzen für die Wissenschaft aus diesen Daten zu ziehen, ist der effiziente und flexible Zugang für Forschende entscheidend. Gleichzeitig hat der notwendige Datenschutz einen hohen Stellenwert beim Zugang zu Forschungsdaten (Mika und Werhan 2025). Besonders dann, wenn diese auf Sozialdaten basieren. Aus diesem Grund steigt die Relevanz der Datenfernverarbeitung für Forschungsdatenzentren (FDZ). Durch Datenfernverarbeitung beziehungsweise den Remote Access können Forschende jederzeit und von jedem Ort auf die benötigten Forschungsdaten zugreifen. Dies erhöht nicht nur die Flexibilität, sondern spart auch wertvolle Zeit, da der physische Zugang zum FDZ entfällt. Besonders in einer zunehmend globalisierten Forschungsgemeinschaft fördert der Remote Access die Zusammenarbeit über geografische und institutionelle Grenzen hinweg, wodurch die Forschungsgeschwindigkeit und -qualität gesteigert werden.

Darüber hinaus ermöglicht Remote Access eine ressourcenschonende Nutzung, da keine umfangreiche lokale Infrastruktur erforderlich ist, um große und komplexe Datensätze zu verarbeiten. Gleichzeitig kann der Datenzugang so gestaltet werden, dass strenge Datenschutz- und Compliance-Richtlinien für die Nutzung öffentlicher Verwaltungsdaten eingehalten werden. Nicht zuletzt trägt Remote Access auch zu einer breiteren und inklusiveren Wissenschaft bei. Durch den einfachen Zugang können Forschende weltweit, unabhängig von ihrem Standort oder ihrer institutionellen Ausstattung, auf wertvolle Daten zugreifen und diese für ihre Arbeiten nutzen. Dies stärkt die Prinzipien von Open Science und fördert eine transparente, reproduzierbare Forschung.

Insgesamt bietet Remote Access eine moderne, flexible und sichere Lösung, die den Anforderungen der heutigen Forschung gerecht wird und gleichzeitig den Zugang zu öffentlichen Verwaltungsdaten erleichtert und verbessert.

Ziel dieses Beitrags ist es zunächst, begriffliche Grundlagen rund um das Thema „Remote Access“ zu klären und Möglichkeiten der technischen Umsetzung aufzuzeigen. Anhand von Beispielen wird gezeigt, wie verschiedene datengebende Institutionen den Zugang zu sensiblen Forschungsdaten gestalten und wo sich das Forschungsdatenzentrum der Rentenversicherung (FDZ-RV) in diesem Feld einordnen lässt.

2. Theoretische Grundlagen

Eine gemeinsame Sprache zwischen Experten und Expertinnen und den IT-Bereichen von Unternehmen zu finden, ist nicht immer leicht. Kim Miriam Dörr und Sandra Aengenheyster (Dörr und Aengenheyster 2019, S. 7) weisen darauf hin, dass IT-Fachbegriffe für fachferne Mitarbeitende schnell abschreckend wirken können:

> „Die Kernbotschaften müssen so übersetzt werden, dass sie von meinen Zielgruppen außerhalb der IT verstanden und angenommen werden."

Das Zitat macht deutlich, dass bereits die Vorstellung zu technischen Begriffen verschieden sein kann und so die Gefahr besteht, dass es zu Missverständnissen zwischen der IT und verschiedenen Stakeholdern (hier die Datengebenden und die Forschenden) kommt. Der folgende Abschnitt soll daher wichtige Begriffe klären, die im Folgenden verwendet werden.

2.1 Abgrenzung technischer Begriffe

Das Vorgehen in diesem Abschnitt folgt dem deduktiven Prinzip „vom Großen zum Kleinen". Dabei wird zwischen der Arbeit mit Forschungsdaten innerhalb der datengebenden Institution und der Arbeit von außerhalb unterschieden.

2.1.1 On-Site und Off-Site

Ob die Daten On-Site oder Off-Site genutzt werden, hängt maßgeblich damit zusammen, wo die Verarbeitung der Daten stattfindet. Eine On-Site-Nutzung bezeichnet die Verwendung von Daten, wenn die Arbeitsschritte zur Analyse der Forschungsdaten innerhalb der Infrastruktur der datengebenden Stelle verarbeitet werden und es den Datengebenden möglich ist, die Ergebnisse, den Output, zu prüfen. Wichtig ist, dass weder Input noch Output die Infrastruktur verlassen, außer über einen von Mitarbeitenden kontrollierten Weg.
Eine gängige Methode ist die Arbeit an einem Gastwissenschafts-Arbeitsplatz oder mithilfe der Datenfernverarbeitung. Beim Gastwissenschafts-Arbeitsplatz lässt sich weiterhin unterscheiden, ob dieser in den Räumlichkeiten der datengebenden Stelle angesiedelt ist oder im Safe-Room eines anderen FDZ („FDZ im FDZ").
Die Off-Site-Nutzung umfasst alle anderen Formen und ist daher wesentlich unschärfer in seiner Definition. Grundsätzlich erfolgt bei dieser Nutzungsform die Bearbeitung der Daten in jedem Fall außerhalb der Infrastruktur der datenliefernden Stelle, während bei einer Datenfernverarbeitung beziehungsweise einem Remote Access die Daten innerhalb der Institution verbleiben.

2.1.2 Remote Access

Während On-Site und Off-Site hauptsächlich den Verarbeitungs- und Speicherort der Daten beschreiben, bezieht sich Remote Access auf den Fernzugriff von IT-Infrastrukturen. Dies kann einzelne Applikationen oder komplette Arbeitsumgebungen umfassen. Remote Access ist ein Oberbegriff für viele Verfahren und Lösungen, die meist über Software umgesetzt werden, wie zum Beispiel der Zugriff auf Netzwerkressourcen via VPN[1]-Tunnel oder der Fernzugriff auf einen Desktop.
Remote Access bietet den Vorteil, dass Forschende von jedem Ort aus mit den Daten arbeiten können. Die datengebende Institution stellt dafür eine IT-Infrastruk-

1 Virtual Private Network (VPN) – Einwahl in eine ferne Netzwerkinfrastruktur über das Internet oder ein Kommunikationsnetz.

tur inklusive Hard- und Software bereit, auf die aus der Ferne zugegriffen werden kann. Gegenwärtig gibt es zwei Hauptmethoden für diesen Zugriff: das Remote-Execution-Verfahren sowie den Remote Desktop.

2.1.3 Remote Execution – Batch Machine

Die Batch Machine und das Remote-Execution-Verfahren sind seit Langem genutzte Remote-Access-Methoden. Einige FDZ, wie die Luxembourg Income Study (LIS) mit dem Tool LISSY, setzen grundlegende Formen dieser Datenfernverarbeitung via E-Mail bereits seit den späten 1980er-Jahren ein (Rowland 2003).
Die Forschenden haben dabei keinen direkten Datenzugriff, können jedoch Programme erstellen und an die datengebenden Institutionen senden. Diese Institutionen führen die Programme anschließend aus und übermitteln die Ergebnisse zurück an die Forschenden. Seit 2005 bietet das sozio-oekonomische Panel (SOEP) am Deutschen Institut für Wirtschaftsforschung (DIW) mit SOEPRemote eine automatisierte E-Mail-Code-Ausführung in Zusammenarbeit mit der LIS an (Goebel 2014). Auch die Deutsche Rentenversicherung Bund bietet Datennutzenden seit 2007 eine Nutzung von Forschungsdaten über Remote Execution[2] an. Forschende können bei der Remote Execution die Daten nicht direkt einsehen, was für die Institution einen geringen Implementierungsaufwand und höheren Schutz der Daten bedeutet.
Der Begriff Remote Execution dient als Beispiel für eine Bezeichnung, die je nach Kontext unterschiedliche Bedeutungen hat. Während in der sozialwissenschaftlichen Forschung die zuvor genannte Definition geläufig ist, bezieht sich Remote Execution oder Remote Code Execution in der IT auf eine Schwachstelle in Computersystemen (TuxCare PR Team 2023).

2.1.4 Remote Desktop

Remote Desktop ermöglicht im Vergleich zum Remote-Execution-Verfahren einen umfassenderen Zugriff auf die Infrastruktur der datenliefernden Institution und ist somit deutlich flexibler.
Remote Desktop umfasst Softwarelösungen, die unterschiedliche Techniken integrieren. Typischerweise werden Methoden verwendet, bei denen mithilfe einer VPN[1]-Software ein VPN-Tunnel zum Netzwerk der IT-Infrastruktur der Institution aufgebaut wird. Dabei kommen im Wesentlichen zwei Verfahren zur Anwendung: Das eine ist VPN, das durch zusätzliche Software auf dem Computer der Forschenden realisiert wird. Das andere ist SSL-VPN, das oft als browserbasierte Softwarelösung fungiert und nur ein Plug-in[3] benötigt. Neben diesen Softwareoptionen gibt es auch Lösungen mit spezieller Hardware.
Remote Desktop kann sich auf Unix- oder Windows-Systeme beziehen und bietet nach Authentifizierung Zugang zu einer Arbeitsumgebung. Nutzerinnen und Nutzer erhalten direktes Feedback, sehen Daten ein und können verschiedene Softwarepakete parallel nutzen und wechseln. Bei Remote Desktop und Remote Execution werden IT-Infrastruktur und Softwarelizenzen von der datengebenden Stelle bereitgestellt.

2.1.5 Safe-Room, FDZ im FDZ

Der Begriff Safe-Room bezeichnet einen gesicherten Arbeitsbereich, ist jedoch eher in Forschungsinstitutionen als in der IT verbreitet. Dort werden gesicherte Arbeitsplätze mit Zugangskontrollen angeboten, während bei Remote-Desktop- und Remote-Execution-Verfahren der Endpunkt nicht kontrolliert wird. In einem Safe-Room können Forschende schwächer anonymisierte Daten

2 https://www.eservice-drv.de/FernrechnenWeb/index.jsp (aufgerufen am 12.09.2025).
3 Plug-in – eine zusätzliche Softwarekomponente, die die Funktion des Browsers erweitert.

verwenden, da die Person, technische Geräte und die Workstation überwacht werden. Das FDZ-RV nutzt diese Infrastruktur, um an verschiedenen Standorten in Deutschland den Zugriff auf Daten zu ermöglichen. Dabei kann die datengebende Stelle die Hardware bestimmen, und die Einbindung erfolgt über einen festen Knotenpunkt. Seit Juni 2023 gibt es eine Vereinbarung zwischen dem GESIS – Leibniz-Institut für Sozialwissenschaften e. V. in Mannheim und FDZ-RV in Berlin über den Safe-Room-Remote-Desktop-Access. Dabei wird ein SINA-PC (siehe Abschnitt 2.2) der Deutschen Rentenversicherung Bund in den Räumlichkeiten von GESIS installiert, um den Zugriff auf die Infrastruktur des FDZ-RV zu ermöglichen. Weitere Standorte für Safe-Rooms des FDZ-RV sind in Köln und München.
FDZ im FDZ ist eine spezielle Version der Safe-Room-Idee. Institutionen bieten ihre Infrastruktur an, sodass Forschende nicht zur datenbereitstellenden Institution reisen müssen, sondern in einem nahegelegenen Institut arbeiten können. Im Unterschied zum Safe-Room kann beim FDZ im FDZ ein Netzwerk aus Forschungsdatenzentren entstehen, das wechselseitig Safe-Rooms für die Datenarbeit bereitstellt. Der Arbeitsplatz im Safe-Room wird hierbei von der betreibenden Institution gestellt, nicht von der datengebenden.

2.1.6 Gastwissenschafts-Arbeitsplatz

Ein Gastwissenschafts-Arbeitsplatz bietet die Möglichkeit, dass Forschende vor Ort in den Räumlichkeiten der datengebenden Stellen arbeiten und direkt deren Infrastruktur nutzen. Hierbei gibt es festgeschriebene Regeln, die der Einhaltung des Datenschutzes dienen (zum Beispiel keine Nutzung eigener Technik am Gastwissenschafts-Arbeitsplatz) und zu deren Einhaltung sich die Nutzenden schriftlich verpflichten. Einschränkungen wie die Netzwerkgeschwindigkeit und Serverkapazität sind minimal. Die Zusammenarbeit mit den Mitarbeitenden der datengebenden Stelle ist eng und sensible Forschungsdaten sind nur schwächer anonymisiert. Das FDZ-RV bietet derzeit sechs solcher Arbeitsplätze in den Räumlichkeiten der Deutschen Rentenversicherung Bund am Standort Berlin an sowie einen Arbeitsplatz in Würzburg.

2.2 Technische Umsetzungen

2.2.1 Off-Site-Nutzung

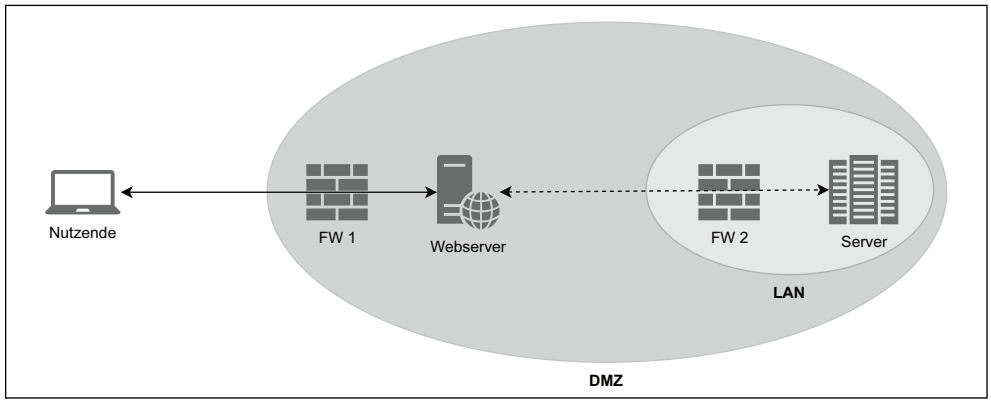
Bei der Off-Site-Nutzung werden verschiedene Arten von Datenprodukten[4] nach einer Datenschutzprüfung zur Nutzung in der Infrastruktur der Datennutzenden zur Verfügung gestellt. Dies erfordert eine sichere Übermittlung an Forschende, meist durch eine web- und cloudbasiert verschlüsselte Übertragung. Die Dateien werden verschlüsselt hochgeladen und können von Forschenden mit einem Passwort entschlüsselt werden. Alternativ erfolgt der Versand über verschlüsselte Datenträger, oder die Campus Use Files und Public Use Files werden direkt auf den Websites der Institutionen angeboten, oft mit Registrierung.

2.2.2 Remote-Execution-Verfahren

Es gibt zwei Ansätze für ein Remote-Execution-Verfahren: einen automatisierten und einen manuellen. In beiden Fällen erstellen die Forschenden eine Programmschrift (Syntax), die dann in der IT-Infrastruktur der datengebenden Institution ausgeführt wird. Das Ergebnis dieser Ausführung wird nach einer datenschutzrechtlichen Überprüfung an die Forschenden zurückgesendet. Beim manuellen Verfahren müssen Mitarbeitende des Instituts die Syntax ausführen. Da dieses Verfahren selten genutzt wird, wird es hier nicht weiter behandelt.

4 Der Beitrag von *Mika* und *Werhan* in dieser Ausgabe nimmt eine systematische Einordnung der Datenproduktarten Scientific Use File (SUF), Public Use File (PUF) und Campus Use File (CUF) vor.

Abbildung 1: Remote Execution mit demilitarisierter Zone



Quelle: eigene Darstellung.

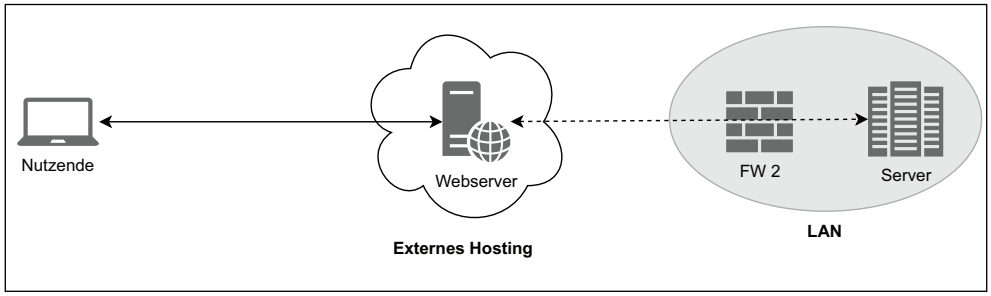
Das automatisierte Remote-Execution-Verfahren setzt grundlegend auf einen Webserver und einen Applikationsserver. Andere Maschinen sind denkbar, werden jedoch der Übersicht halber nicht betrachtet. Forschende interagieren ausschließlich mit einem Onlineportal, das auf dem Webserver installiert ist. Dieser Server kann entweder vom Institut selbst betrieben werden und in der sogenannten demilitarisierten Zone[5] (DMZ) des Netzwerks platziert sein oder von einem externen Anbieter außerhalb des eigenen Netzwerks gehostet werden. Aus diesem Onlineportal werden die Syntaxen abgeholt und in das interne Netzwerk kopiert. Da es sich um Plain-Text-Daten[6] handelt, können diese problemlos über das https-Protokoll übertragen werden. Das Onlineportal fungiert somit passiv als Austauschplattform. Im Netzwerk der datengebenden Stelle wird der Code überprüft und ausgeführt, häufig in einer separierten Netzwerkzone. Nach der Ausführung werden die Ergebnisse, meist als Log-Dateien der Statistiksoftware, zurück an das Onlineportal gesendet, wo sie von den Forschenden heruntergeladen werden können.
Ein mögliches Konfigurationsbeispiel ist in Abbildung 1 dargestellt. Die Mikrodaten verbleiben auf dem Server der datengebenden Einrichtung und werden nicht transferiert. Die Syntax wird über eine SSL[7]-geschützte Webapplikation heruntergeladen. Dies erfordert Anpassungen des Sicherheitskonzepts wie etwa Änderungen an der Firewall, da ein Webserver innerhalb der DMZ betrieben werden muss. Die Rechner im LAN[8] benötigen dann Zugriffsberechtigungen für diesen Webserver. Abbildung 2 stellt ein mögliches Szenario ohne eine separate DMZ dar. In diesem Fall sind keine zusätzlichen Anpassungen der bestehenden IT-Infrastruktur erforderlich, da der Webserver außerhalb des eigenen Netzwerks installiert ist.

2.2.3 Beispiel für ein Remote-Execution-Verfahren

Die webbasierte Job Submission Application (JoSuA) ist ein Softwareprodukt des Instituts zur Zukunft der Arbeit (IZA). In Deutschland

5 Demilitarisierte Zone – ein Bereich im Netzwerk, welcher für das Hosting von Webservices Ports offenhält. Dieser Bereich ist durch eine Firewall vom internen Netzwerk getrennt.
6 Plain-Text-Daten – Daten, die ohne Formatierung als einfacherer Text gespeichert sind.
7 Secure Sockets Layer – Verschlüsselungsprotokoll, das vor allem für Websites genutzt wird (https://...).
8 Local Area Network (LAN) – das lokale, häufig auch lokal beschränkte Netzwerk.

Abbildung 2: Remote Execution ohne demilitarisierte Zone



Quelle: eigene Darstellung.

wird JoSuA unter anderem vom Bundesinstitut für Berufsbildung (BiBB), dem Institut für Arbeitsmarkt- und Berufsforschung (IAB) und dem Institut zur Qualitätsentwicklung im Bildungswesen (IQB) eingesetzt. JoSuA stellt eine Remote-Execution-Lösung dar, die den Datenverarbeitungsprozess stark automatisieren kann. Die Konfiguration von JoSuA beinhaltet ein Webportal, das sowohl Eingaben der Nutzenden als auch der Mitarbeitenden entgegennimmt und weiterleitet, sowie Server, auf denen die Datenauswertungen stattfinden.

2.2.4 Frontend und Backend

Das Herzstück von JoSuA ist ein Internetportal, das für die Datennutzenden bereitgestellt wird. Forschende erhalten einen Account und können sich damit im Portal anmelden. Auf einem Dashboard werden ihnen ihre eigenen Projekte präsentiert. Innerhalb dieses Dashboards haben sie die Möglichkeit, Auswertungen durchzuführen. Diese Auswertungen können nur auf den Datenbeständen durchgeführt werden, die für das jeweilige Projekt freigegeben sind. Zusätzlich zum Webportal ist der Erwerb von Statistiksoftware für den Betrieb von JoSuA erforderlich. JoSuA ist mit verschiedenen Produkten wie STATA, SPSS, Python, R oder MATLAB kompatibel.

Im Backend wird Forschenden die Rolle „User“ zugewiesen. Mitarbeitende der datengebenden Stellen erhalten die Rolle „Operator“. In dieser Position haben Operator nahezu dieselbe Ansicht auf das Portal wie User, mit wenigen Ausnahmen. Diese zusätzlichen Funktionen betreffen hauptsächlich den Datenschutz. Über die Operator-Ansicht können Mitarbeitende der datengebenden Stellen die Outputs abgeschlossener Aufgaben für Forschende freigeben und Forschende sowie Software den Projekten zuordnen.
Die dritte Rolle ist die des „Administrator“. Diese sollte idealerweise nicht von jemandem übernommen werden, der bereits als Operator tätig ist. Der Administrator hat eine ähnlich gestaltete Portalansicht wie die anderen Rollen. Darüber hinaus kann der Administrator technische Einstellungen der JoSuA-Installation konfigurieren.

2.2.5 Workflow

Den Forschenden stehen im Frontend zwei Modi für die Auswertungen zur Verfügung: ein interner Modus und ein Veröffentlichungsmodus. Nach dem Hochladen der Syntax und der Auswahl der Forschungsdaten können sie den gewünschten Modus wählen. Im internen Modus werden die Auswertungen automatisch an den Server

gesendet. Nach Ausführung des Skripts wird die Log-Datei dem Nutzenden angezeigt, jedoch mit einem Wasserzeichen versehen, um eine Weiterverarbeitung zu verhindern. Dieses Verfahren ermöglicht auch die Darstellung von Grafiken, die ebenfalls durch Wasserzeichen geschützt sind.
Abhängig von der Konfiguration der JoSuA-Installation können Forschende im internen Modus frei experimentieren. Sollen Auswertungen oder Grafiken veröffentlicht werden, müssen diese in den Veröffentlichungsmodus übertragen werden. In diesem Fall erhalten die Operators eine Benachrichtigung über eine angeforderte Freigabe und müssen darauf reagieren. Wenn die Auswertungen als sicher eingestuft werden, erfolgt die Freigabe durch die Operators, sodass diese dann aus dem Webportal heruntergeladen werden können.

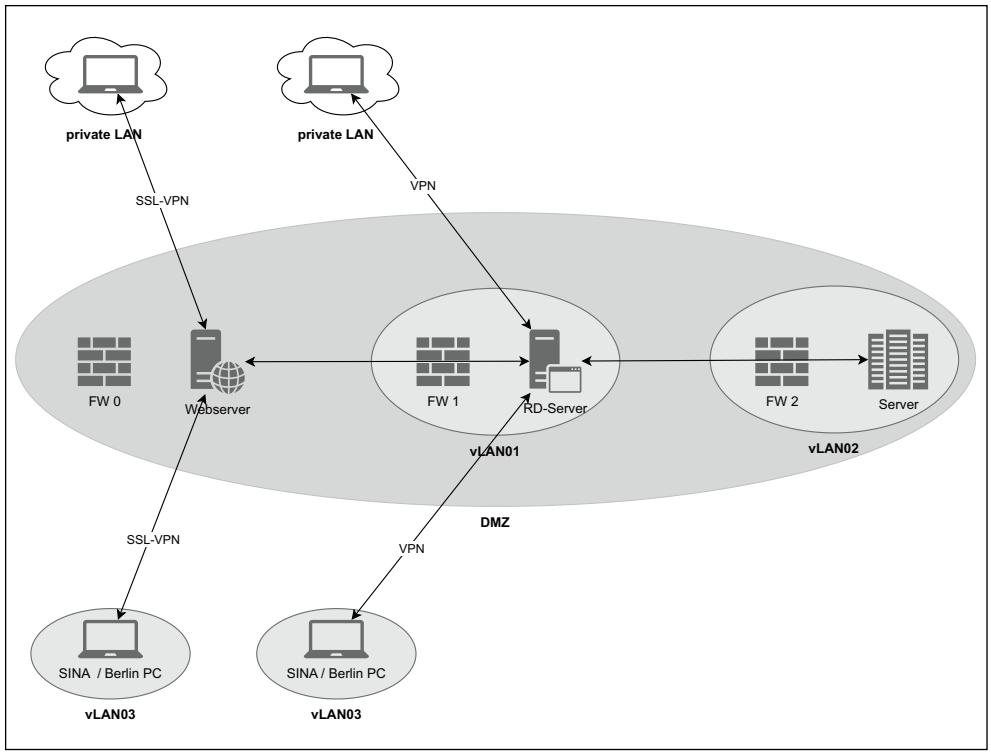
2.2.6 Remote-Desktop-Verfahren

Die technischen Anforderungen für die Remote-Desktop-Lösung sind höher. Nutzende müssen auf Geheimhaltung achten. Für die Nutzung von Remote Desktops gelten hinsichtlich der Geheimhaltung die allgemeinen Regeln der Telearbeit. Das Bundesamt für Sicherheit in der Informationstechnik (BSI) hat Empfehlungen veröffentlicht, um eine sichere Arbeitsumgebung zu gewährleisten (Bundesamt für Sicherheit in der Informationstechnik – BSI 2021). Das BSI betont, dass die IT-Umgebung, die von den Nutzenden verwendet wird, berücksichtigt werden muss. Hier ist zu beachten, ob Forschende ihre eigene IT-Infrastruktur, wie zum Beispiel ihren privaten Computer, nutzen. Ebenso ist das Netzwerk von Bedeutung, in dem sich die Forschenden bewegen. Hier wird zwischen privaten und öffentlichen sowie institutionellen Netzwerken unterschieden. Neben der Geheimhaltung ist bei der Remote-Desktop-Lösung eine eindeutige Identifizierung erforderlich, idealerweise durch ein zweistufiges Verfahren mit biometrischen Merkmalen.

2.2.7 Beispiele für Remote-Desktop-Verfahren mit zusätzlicher Hardware

Als Beispiele sollen hier zwei konkrete Lösungen vorgestellt werden. Eine davon ist der BerlinPC, den das IT-Dienstleistungszentrum Berlin (ITDZ Berlin) entwickelt hat. Diese Lösung liefert ein Betriebssystem auf einem speziell vorbereiteten USB-Stick aus, der das Hostsystem der Nutzenden in einer sicheren, Linux-basierten Umgebung bootet. Innerhalb dieser Umgebung kann eine VPN-Verbindung zu einem Terminalserver hergestellt werden, auf dem dann gearbeitet wird. Für diese Lösung sind außer dem USB-Stick und einem privaten Computer keine zusätzlichen Hardware-Komponenten erforderlich.
Eine weitere Lösung, die in der deutschen Verwaltung eingesetzt wird, zum Beispiel in der Deutschen Rentenversicherung Bund, stammt von der Firma secunet Security Networks AG aus Köln in Zusammenarbeit mit dem BSI. Die Sichere Inter-Netzwerk-Architektur (SINA) zielt darauf ab, eine sichere Arbeitsumgebung in unsicheren Netzwerken zu schaffen. Hierbei werden Endgeräte mit einem speziellen Linux-Betriebssystem ausgestattet, in dem zusätzliche Betriebssysteme wie Windows in einer virtuellen Umgebung gebootet werden können. Durch dieses Setup ist „eine gesicherte Kommunikation über bestehende, nicht vertrauenswürdige Verbindungen“ (Bundesamt für Sicherheit in der Informationstechnik – BSI 2016, S. 12) möglich. Diese Lösung erfordert jedoch speziell angepasste Hardware, die sogenannten SINA-Workstations. Das BSI gewährleistet dabei eine Vertrauenswürdigkeit bis zur Stufe „geheim“.
Beide Lösungen bieten eine sichere Umgebung auf der Seite der Nutzenden und werden unter den Begriffen „auto vpn“ oder „mobile vpn“ zusammengefasst. In dieser geschützten Umgebung wird eine VPN-Verbindung mit dem Netzwerk der datengebenden Stelle aufgebaut. Dank ihres doppelten

Abbildung 3: Remote-Desktop-Verfahren



Quelle: eigene Darstellung.

Schutzes – gehärtetes Betriebssystem[9] und VPN-Verbindung zur IT-Infrastruktur – sind diese Systeme sehr sicher.

2.2.8 Beispiele für Remote-Desktop-Verfahren ohne zusätzliche Hardware

Alternativ gibt es Lösungen, bei denen eine VPN-Verbindung aufgebaut wird, ohne dass zusätzliche Anforderungen an die IT-Umgebung des Hostsystems gestellt werden. Ein Beispiel dafür ist die Humboldt-Universität zu Berlin, die ihren Studierenden Zugang zur IT-Infrastruktur der Universität über einen VPN-Tunnel gewährt, der mittels SSL-VPN oder OpenVPN realisiert wird.

Eine weitere Möglichkeit, einen Remote-Desktop-Arbeitsplatz anzubieten, besteht darin, den Remote-Desktop direkt im Browser anzuzeigen. Der Remote-Desktop-Client ist in diesem Fall eine Webanwendung, deren wesentliche Anforderung an die IT-Umgebung lediglich ein aktueller Browser ist, der das Speichern von Cookies zulässt. Das FDZ des Leibniz-Instituts für Bildungsverläufe (FDZ-LIfBi) bietet mit Remote-NEPS eine solche Lösung an.

Abbildung 3 stellt die beschriebenen Remote-Desktop-Verfahren dar. In dieser Abbildung sind Lösungen dargestellt, die am

9 Ein Betriebssystem mit erhöhter Sicherheit, welches besser vor Angriffen geschützt ist.

eigenen Computer genutzt werden können, und solche, die mit zusätzlicher Hardware realisiert werden. Für die Lösungen mit zusätzlicher Hardware wurde zur besseren Lesbarkeit das virtuelle Netzwerk vLAN03 in das Bild integriert. Tatsächliche Umsetzungen dieser Technik sind wesentlich komplexer, jedoch bietet die Abbildung eine vereinfachte Darstellung zur Verdeutlichung des Prinzips.

Für ein Safe-Room-Setup oder ein FDZ im FDZ wird außerhalb der Räumlichkeiten der datengebenden Stelle ein eigener Arbeitsplatz eingerichtet. Dieser Arbeitsplatz wird in einer vertrauenswürdigen Umgebung eingerichtet, wie beispielsweise an Universitäten oder in Verwaltungseinrichtungen in Deutschland. Ein Raum mit Zugangsbeschränkung wird so ausgestattet, dass eine sichere Verbindung zur datengebenden Stelle gewährleistet werden kann. Dabei ist die IT-Seite des Safe-Rooms weitgehend bekannt und die Sicherheit des Netzwerks der Einrichtung abschätzbar und überprüfbar. Die Verbindung zur datenliefernden Institution erfolgt über eine VPN-Verbindung, die zusätzlich durch eindeutige Kennungen des Fernrechnen-PCs gesichert wird. Zu diesen Kennungen können die IP-Adresse des Computers beziehungsweise der Institution oder die Mac-Adresse des Computers gehören. Eine feste IP-Adresse ist oft eine Anforderung, die nicht immer erfüllt werden kann, da insbesondere private Internetzugänge in Deutschland häufig mit wechselnder IP betrieben werden.

3. Internationale Beispiele

Bereits im Jahr 2012 wurde vom inzwischen beendeten Projekt „Data without Boundaries (DwB)“ ein Paper veröffentlicht, in dem verschiedene Remote-Access-Zugänge zu Mikrodaten der öffentlichen Verwaltung in Europa verglichen wurden. Die Autoren arbeiten zwei Aspekte heraus: Bei vielen Gemeinsamkeiten der Lösungen gibt es eine große Varianz, die einerseits für datengebende Stellen viele Möglichkeiten bieten, Ideen für eigene Lösungen zu sammeln, andererseits aber eine Herausforderung auf dem Weg zu einer vereinheitlichten Datenstrategie in Europa (Gürke et al. 2012, S. 24) darstellen. Die im Folgenden beschriebenen Beispiele wurden gewählt, weil sich häufig auf sie bezogen wird beziehungsweise weil sie direkt an Lösungen für deutsche Institutionen beteiligt waren.

3.1 Statistics Sweden

Statistics Sweden betreibt seit dem Jahr 2005 eine Remote-Desktop-Lösung namens MONA[10] (Microdata Online Access). Das MONA-System bietet einen sicheren Zugang zu den Mikrodaten von Statistics Sweden über das Internet. Für die Arbeit mit den Daten steht eine Vielzahl an (Statistik-)Softwareprodukten zur Verfügung. MONA bietet einen Windows Desktop. Aus Sicherheitsgründen können die Forschenden Daten zwischen ihrem eigenen Computer und MONA nur über eine webbasierte Schnittstelle austauschen. Die Forschenden haben jeweils einen eigenen Speicherplatz für Ergebnisdateien, Codebibliotheken und andere vom Benutzer erstellte Dateien. Um diese Dateien von MONA herunterzuladen, wird ebenfalls ein Webinterface verwendet, über das die Dateien automatisch an die Forschenden gesendet werden.

3.2 UK Data Service

Im Jahr 2008 führte das britische Datenarchiv ein Pilotprojekt mit der Bezeichnung Secure Data Service[11] (SDS) ein. Der

10 https://www.scb.se/en/services/ordering-data-and-statistics/ordering-microdata/mona--statistics-swedens-platform-for-access-to-microdata/about-mona/ (aufgerufen am 12.09.2025).

11 https://www.data-archive.ac.uk/about/grants-and-projects/research-infrastructure/secure-data-service/ (aufgerufen am 12.09.2025).

Hauptzweck war die Erleichterung des sicheren Forschungszugangs zu detaillierten Mikrodaten, die sonst nur vor Ort bei den Datenproduzenten – oder gar nicht – verfügbar gewesen wären. 2011 führte der SDS ein Remote-Access-System, das Secure Lab, ein, das auf einer Citrix-Infrastruktur basiert. Es verwandelt den Computer der Forschenden in ein „Remote-Terminal“, das Zugang zu Daten, Statistiksoftware und auf einen zentralen, sicheren Server des britischen Datenarchivs hat. Dieser Zugang ist nicht für beliebige Computer als Endpunkte, sondern für einen fest gewählten Computer konzipiert.
Das System ist insofern flexibel, als der Zugang je nach den Wünschen der Datenverwalter auf bestimmte Benutzerinnen und Benutzer (sichere Personen) und/oder bestimmte Orte (sichere Räume/Maschinen) beschränkt werden kann. Die Arbeit findet nach wie vor mit einer vollständigen grafischen Windows-Benutzeroberfläche statt. Die Forschenden benötigen einen Webbrowser und ein Plug-in. Die gesamte Datenverarbeitung findet auf einem sicheren Server statt, der alle Anfragen zentral verarbeitet und Informationen über die Ergebnisse zurückgibt. Es folgt dem Prinzip, dass die Auswertungen zu den Daten und nicht die Daten zur Auswertung gebracht werden. Die statistischen Ergebnisse werden von SDS-Mitarbeitenden für die statistische Offenlegungskontrolle innerhalb des zentralen Servers geprüft und, wenn sie als unbedenklich eingestuft werden, per E-Mail an den Forschenden weitergeleitet. Die gesamte Sitzung wird als Screen Record, ein Mitschnitt des Bildschirminhaltes, aufgezeichnet und archiviert.

3.3 Statistics Netherlands (CBS)

CBS bietet seit 2006 ein spezielles IT-Netz mit Terminalserver (CITRIX), das den Zugriff auf Mikrodaten über eine sichere Internetverbindung vom eigenen Arbeitsplatz der Forschenden ermöglicht, an. Das Netzwerk ist in drei verschiedene Zonen unterteilt, nämlich den Remote-Access-Webserver, die CITRIX-Farm (Domain Controller, Fileserver und so weiter) und die statistische Produktionszone. Diese dritte Zone ist für Nicht-CBS-Forschende nicht zugänglich.
Der Remote-Access-Zugang wird über eine SSL-VPN Lösung via CITRIX ermöglicht, die auf der Workstation der Forschenden am Institut der Forschenden eingerichtet wird. Auf dieser Workstation besteht die Möglichkeit, nach erfolgreicher Anmeldung eine der bereitgestellten (Statistik-)Anwendungen zu nutzen. Um den Remote-Access zu nutzen, starten die Forschenden den Browser, öffnen die Website von Remote Access Microdata und melden sich mit Passwort und einem Token an. Verfügen die Forschenden über mehr als ein aktives Forschungsprojekt und damit eine entsprechende Anzahl von Benutzerkonten, erscheint über ein Pop-up-Fenster die Aufforderung, eines dieser Konten für die Anmeldung auszuwählen. Für jede Remote-Acccess-Instanz werden ein Export- und ein Importverzeichnis eingerichtet. Das Importverzeichnis wird verwendet für Syntaxen oder Codelisten, die im Rahmen des Forschungsprojekts verwendet werden. Hierfür gibt es ein webbasiertes Portal, welches für den Datenupload bereitsteht. Der Exportordner wird verwendet, um Output an die RDC-Mitarbeitenden zur Überprüfung zu senden.

3.4 LIS Cross National Data Center

Heute ist von der Luxembourg Income Study nur noch das Akronym LIS geblieben (LIS: Cross-National Data Center in Luxembourg 2024).
Das in Abschnitt 2.1 bereits erwähnte Tool LISSY wurde seit seiner Einführung in den 1980er-Jahren deutlich weiterentwickelt. Heute ist es ein Remote-Execution-System, das es Forschenden ermöglicht, von einem entfernten Standort aus auf die angebotenen Mikrodaten zuzugreifen und dabei die von den datenliefernden Ländern geforderten Datenschutzbestimmungen einzuhalten. Es

bietet einen sicheren Zugang zu den Mikrodaten über eine webbasierte Schnittstelle für die Übermittlung von Aufträgen.

4. Fazit und Empfehlung

Die vorangegangenen Ausführungen verdeutlichen das Interesse in Deutschland, Forschungsdaten der wissenschaftlichen Forschung zugänglich zu machen. Viele Institutionen bieten Forschungsdaten an, die selbst Forschungsdaten erstellen oder sich als Forschungsinfrastrukturen verstehen. Diese Infrastrukturen sind in der deutschen Forschungsdatenlandschaft besonders wichtig, da sie aktuell als Innovationsmotoren gelten. Beispielhaft ist das Research Data Center of IZA (IDSC) zu nennen, das mit JoSuA eine Möglichkeit geschaffen hat, Forschungsdaten per Remote Access zugänglich zu machen. Auch das Deutsche Zentrum für Hochschul- und Wissenschaftsforschung (DZHW) bietet mit seiner Remote-Desktop-Lösung hauptsächlich Forschungsdateninfrastruktur.

Im internationalen Vergleich könnten deutsche Institute bei der Umsetzung von Remote-Access-Maßnahmen etwas schneller sein. In den letzten Jahren hat sich jedoch, auch durch die Unterstützung der Bundesregierung und der Nationalen Forschungsdateninfrastruktur (NFDI) eine Aufbruchstimmung entwickelt. Bestehende Netzwerke werden neu belebt und es werden neue Netzwerke gegründet. Die Pandemieeinschränkungen haben zusätzlich verdeutlicht, wie notwendig Remote-Access-Technologien sind.

Für Institutionen, die derzeit keinen Remote Access zu ihren Forschungsdaten anbieten, gibt es zwei mögliche Ansätze: Remote Execution oder Remote Desktop. Aus Sicht der Forschenden wird eine Remote-Desktop-Lösung bevorzugt, da sie direktes Feedback bietet und so das Arbeiten sowie die Fehlerbehebung beim Schreiben der Syntax vereinfacht. Ein weiterer Vorteil ist, dass auf dem Remote Desktop nicht nur statistische Software installiert werden kann, sondern auch Textverarbeitungsprogramme, Versionskontrollsysteme und mehr. Diese Vielseitigkeit erleichtert den Forschenden die Weiterverarbeitung ihrer Analysen in derselben Umgebung und ermöglicht kollaboratives Arbeiten, wodurch Synchronisationsprobleme seltener auftreten.

Auch für Datengeber ist der Remote Desktop attraktiv, da dadurch interne und externe Arbeitsplätze besser integriert werden können. Lösungen mit Computern, die über eine SINA verfügen, sind ebenfalls sinnvoll, weil sie eine eindeutige Sicherheit des Endgeräts bieten. Allerdings bleiben bei der Nutzung von Remote-Desktop-Systemen rechtliche Fragen offen. In Deutschland gab es in den letzten zwei Jahren einen klaren Trend zum Homeoffice, was zur Beschaffung vieler Heimarbeitsplatzsysteme führte. Vertrauen spielt hier eine große Rolle; im Forschungsbereich scheint dies jedoch geringer zu sein, da oft nach Maßnahmen zum Schutz vor Bildschirmfotos gefragt wird – ein Aspekt, der im Homeoffice weniger Beachtung findet.

Neben den Remote-Desktop- gibt es auch Remote-Execution-Verfahren. Obwohl diese oft als veraltet gelten, bieten moderne Webanwendungen wie LISSY oder JoSuA großen Komfort, beispielsweise direktes Feedback zu Syntaxskripten und den Zugriff auf Forschungsergebnisse von beliebigen Arbeitsplätzen aus. Dennoch empfinden viele Forschende die Trennung von den eigentlichen Daten als unangenehm. Die hohe Sicherheit spricht jedoch für Remote-Execution-Verfahren, da sie ohne zusätzliche Löcher in der Firewall auskommen können. So können JoSuA-Webserver außerhalb der Firewall platziert werden, während die Terminal-/Application-Server und Daten dahinter sicher bleiben. Zudem lassen sich IT-Ressourcen gut planen, da Anfragen an den Application Server in der Reihenfolge abgearbeitet werden können, in der sie ankommen. Dies reduziert den Lizenzbedarf

für Softwareprodukte und macht das System skalier- und planbar.
Ein weiterer Vorteil von Remote-Execution-Verfahren liegt in der leichteren Umsetzung von Geheimhaltung, da auf Mikrodaten kein direkter Zugriff erfolgt und somit keine Möglichkeit besteht, diese beispielsweise abzufotografieren.
Das FDZ-RV bietet derzeit ein Remote-Execution-Tool, sowie einen Remote-Desktop-Zugang am Gastwissenschafts-Arbeitsplatz in den eigenen Räumlichkeiten oder mittels SINA-Workstations, die in sicheren Umgebungen bei anderen Forschungsdatenzentren (FDZ in FDZ) aufgestellt sind, an. Derzeit gibt es keine konkreten Pläne ein Remote-Desktop-Verfahren einzurichten, das Forschende auch von ihren privaten Rechnern aus nutzen können. Angesichts der steigenden Nachfrage kann sich dies in Zukunft jedoch ändern.

Literatur

Bundesamt für Sicherheit in der Informationstechnik – BSI (2016): Sichere Inter-Netzwerke. Bonn: Bundesamt für Sicherheit in der Informationstechnik – BSI.

Bundesamt für Sicherheit in der Informationstechnik – BSI (2021): Tipps für sicheres mobiles Arbeiten. Bonn: Bundesamt für Sicherheit in der Informationstechnik – BSI.

Dörr, Kim Miriam und *Aengenheyster, Sandra* (2019): Einleitung: Grundlagen der IT-Kommunikation. In: Praxishandbuch IT-Kommunikation. Springer Berlin Heidelberg, S. 254.

Goebel, Jan (2014): Job submission instructions for the SOEPremote System at DIW Berlin: Update 2014. SOEP Survey Papers 195: Series G.

LIS: Cross-National Data Center in Luxembourg (2024): About LIS. https://www.lisdatacenter.org/about-lis/ (aufgerufen am 09.10.2024).

Mika, Tatjana und *Werhan, Katharina* (2025): Adaptive Datenschutzkonzepte in der Praxis des Forschungsdatenzentrums der Rentenversicherung. In: Deutsche Rentenversicherung 80 (3), S. 209–228.

Rowland, Sandra K. (2003): An Examination of Monitored, Remote Microdata Access Systems. NAS Workshop.

Schiller, David/Gadouche, Kamel/Gürke, Christopher/Bond, Steve und *Alvheim, Atle* (2012): Report on the State of the Art of current SC in Europe. Data without Boundaries.

TuxCare PR Team (2023): Remote code execution attack: what it is, how to protect your systems. tuxcare.com: https://tuxcare.com/blog/remote-code-execution-attack-what-it-is-how-to-protect-your-systems/ (aufgerufen am 14.03.2023).

Anschrift des Verfassers:

Sebastian Ellert
Deutsche Rentenversicherung Bund
Abteilung Forschung und Entwicklung
Forschungsdatenzentrum der
Rentenversicherung (FDZ-RV)
Ruhrstraße 2
10709 Berlin