FORSCHUNGSDATENZENTRUM
der Bundesagentur für Arbeit im Institut für Arbeitsmarkt- und Berufsforschung

Anleitungen für Datenfernverarbeitung und Gastaufenthalte am FDZ der BA im IAB

Inhaltsverzeichnis

1 Einleitung **4**

2 Grundlagen **6**
2.1 Projektverzeichnisse auf den FDZ-Servern . . . 6
2.2 Weitere Ressourcen . . . 7
2.2.1 Ado-Files (Stata) . . . 7
2.2.2 R-Packages . . . 9
2.2.3 Zuspielen externer aggregierter Datensätze . . . 9
2.3 Einrichtung einer Testumgebung am eigenen Arbeitsplatz . . . 10
2.3.1 Einrichtung eines Projektverzeichnisses . . . 10
2.3.2 Download und Vorbereitung der Testdaten . . . 10
2.3.3 Verwendung der FDZ-Templates . . . 11
2.3.4 Programmpakete und externe Aggregatdaten . . . 11
2.3.5 Testen der Skripte vor Gastaufenthalt und Datenfernverarbeitung mit JoSuA . . . 12

3 Gestaltung der Skripte **13**

4 Datennutzung mit JoSuA (Job Submission Application) **14**
4.1 Grundsätzliches zur Nutzung von JoSuA . . . 14
4.2 Internal Use Modus und Presentation / Publication Modus . . . 16
4.3 Ablauf der kontrollierten Datenfernverarbeitung mit JoSuA . . . 17

5 Gastaufenthalte **20**
5.1 Vorbereitungen für den Gastaufenthalt . . . 20
5.2 Während des Gastaufenthalts . . . 21
5.3 Erhalt der geprüften Ergebnisse nach Gastaufenthalt . . . 22

6 Datenfernverarbeitung per Remote-Desktop **23**
6.1 Vorbereitungen für den Remote-Desktop . . . 23
6.2 Arbeiten mit dem Remote-Desktop . . . 24
6.3 Validierung und Export von Ergebnissen . . . 24

7 Ergebnisübermittlung **26**
7.1 Datenschutzrichtlinien . . . 26
7.2 Zulässige Ausgabeformate . . . 26

8 Publikation der Ergebnisse **28**

9 Sonderregelungen der Softwarenutzung **29**

10 Bedienungsanleitung JoSuA **31**
10.1 Login . . . 31
10.2 Start . . . 31
10.3 Home . . . 32
10.4 Projects: Projektübersicht . . . 34
10.5 Submission: Hochladen von Aufträgen . . . 37
10.6 Stoppen von Jobs . . . 38
10.7 Transfers: Übermittlung von Skripten vor oder nach dem Gastaufenthalt . . . 38
10.8 Technische Restriktionen in JoSuA . . . 39
10.9 Hilfe bei Job-Abbrüchen in JoSuA . . . 40
10.10 Fehlerbehebung in JoSuA . . . 41
10.10.1 Dateien können nicht hochgeladen werden . . . 41
10.10.2 Freigabe der Ergebnisse im Internal Use Modus dauert ungewöhnlich lang (Status „censoring") . . . 41
10.10.3 Ergebnisdateien fehlen . . . 41
10.10.4 Fehlermeldung im Log-File "File not found" . . . 42
10.10.5 Log-File master.log / Fehlermeldung "log-file already open" . . . 42
10.10.6 Fehlermeldung "File not ok (suspicious mismatch between file suffix and actual file type)" . . . 42
10.10.7 Fehlermeldung "Error!" . . . 42

11 Bedienungsanleitung Remote-Desktop **43**
11.1 Login Remote-Desktop-Zugang . . . 43
11.2 Zusätzliche Navigationselemente . . . 45
11.3 Trennen oder Abmelden? . . . 47
11.4 Passwort ändern . . . 47
11.4.1 Passwortkonventionen . . . 48
11.4.2 Kontosperre . . . 49

1 Einleitung

Das vorliegende Dokument (im Folgenden „Anleitungen" genannt) ist eine Sammlung unterschiedlicher Anleitungen des FDZ, die Sie bei der Nutzung unserer Datenzugangswege unterstützen sollen. Hier erhalten Sie folgende Informationen:

- Wie sieht das Projektverzeichnis in der IT-Infrastruktur des FDZ aus? (siehe Kapitel 2.1)
- Wie kann ich Programmpakete und externe aggregierte Daten nutzen? (siehe Kapitel 2.2)
- Wie kann ich meine Skripte außerhalb des FDZ testen? (siehe Kapitel 2.3)
- Wie müssen die Skripte gestaltet sein? (siehe Kapitel 3)
- Wie funktioniert die Datenfernverarbeitung mit JoSuA? (siehe Kapitel 4)
- Wie läuft ein Gastaufenthalt an einem der FDZ-Standorte ab? Was muss ich im Vorfeld von Gastaufenthalten beachten? (siehe Kapitel 5)
- Wie läuft die Datenfernverarbeitung per Remote Desktop ab? (siehe Kapitel 6)
- Wie erhalte ich meine Ergebnisse? (siehe Kapitel 7)
- Was muss ich bei der Veröffentlichung von Ergebnissen beachten? (siehe Kapitel 8)
- Was ist zu beachten, wenn ich eine andere Software als Stata nutzen möchte? (siehe Kapitel 9)
- Wie bediene ich JoSuA? (siehe Kapitel 10)
- Was kann ich bei Problemen in JoSuA selbst zur Behebung beitragen? (siehe Kapitel 10.10)
- Wie gelange ich in die Remote Desktop Umgebung und wie bediene ich sie? (siehe Kapitel 11)

Neben diesem Dokument sind zwingend die „**Richtlinien zur Gestaltung von Auswertungsprogrammen und Analyseergebnissen**" (im Folgenden „Richtlinien" genannt) zu beachten. Dies ist das zentrale Dokument, das Ihnen zeigt, welche Vorgaben wir für den Aufbau von Auswertungsprogrammen/Skripten machen und wie Sie datenschutzkonforme Ergebnisse generieren. Bitte nutzen Sie hierzu immer auch die vom FDZ zur Verfügung gestellten **Stata-Templates für Auswertungsprogramme** (im

Folgenden „FDZ-Templates" genannt) für die Arbeitsschritte Datenaufbereitung, Deskription, Analyse und Ergebnisausgabe, in denen beispielhaft die Richtlinien umgesetzt sind.

Zusätzlich stehen Ihnen auf der FDZ-Website für alle Datenprodukte FDZ-Datenreporte, Auszählungen und Testdaten zur Verfügung. Weitere Arbeitshilfen und Publikationen zur Arbeit mit den FDZ-Daten finden Sie auf der FDZ-Website unter „Zentrale Arbeitshilfen".

Bitte nutzen Sie diese Dokumente und Arbeitshilfen nicht nur, bevor Sie mit den Analysen der Daten beginnen, sondern auch als Nachschlagewerk, wenn Probleme auftreten. Bitte bedenken Sie immer, dass das FDZ weit über 1000 Nutzende betreut und auf Ihre Mitwirkung angewiesen ist, um seinen Service in angemessener Zeit aufrecht erhalten zu können. Kontaktieren Sie uns (iab.fdz@iab.de) bei Problemen bitte erst dann, wenn Sie in unseren Dokumenten keine passende Hilfestellung finden können. Bei Problemen mit JoSuA informieren Sie sich bitte immer zuerst auf unserer Website, ob allgemeine Störungen des Dienstes vorliegen oder passende Hilfestellungen in den JoSuA-FAQs zu finden sind.

2 Grundlagen

2.1 Projektverzeichnisse auf den FDZ-Servern

Jedes Forschungsprojekt erhält pro FDZ-Datenprodukt ein eigenes Verzeichnis (**fdzXXXX**). Verschiedene FDZ-Datenprodukte können nicht gemeinsam verwendet oder auf Einzelbeobachtungsebene zusammengeführt werden, auch wenn sie im gleichen Projekt genutzt werden. Falls Sie mehr als ein FDZ-Datenprodukt nutzen, können Sie selbst herausfinden, welches FDZ-Datenprodukt in welchem Projektverzeichnis abgelegt ist. Erstellen Sie sich dafür im Rahmen der Datenfernverarbeitung Übersichtslisten zu den vorliegenden Dateien je Projektverzeichnis. Ein Beispiel finden Sie in den FDZ-Templates (`master.do`).

Jedes Projektverzeichnis enthält die in Tabelle 1 aufgelisteten Unterverzeichnisse. Das Anlegen weiterer Verzeichnisse oder Unterverzeichnisse ist nicht erlaubt.

Tabelle 1: Unterverzeichnisse im Projektverzeichnis

Tabelle 2.1: Subdirectories in the project directory

Name	Beschreibung	Zugriff
orig	In diesem Verzeichnis werden alle beantragten Originaldaten bereitgestellt. Auch ggf. eingesendete Aggregatdaten (siehe Kapitel 2.2.3) befinden sich hier. Für dieses Verzeichnis erhalten Sie nur Leserechte.	Auf Daten in den Verzeichnissen **orig** und **data** können Programme auch über **JoSuA** zugreifen.
data	Hier werden alle generierten Datensätze abgelegt. Pro Verzeichnis sind maximal 30 GB erlaubt. Sollte das Maximum überschritten werden, muss der Speicherbedarf reduziert werden. Ein Beispiel dazu finden Sie in den FDZ-Templates.	Auf Daten in den Verzeichnissen **orig** und **data** können Programme auch über **JoSuA** zugreifen.
prog	In diesem Ordner befinden sich alle Skripte (z.B. Stata-Do-Files, R-Skripte, .m-Files) für den Gastaufenthalt sowie ggf. weitere Dateien (z. B. ado-Dateien, R-Packages).	**Skripte und Ergebnisdateien** werden **nicht** automatisch zwischen **JoSuA** und dem **Projektverzeichnis** abgeglichen.

Name	Beschreibung	Zugriff
log	Dieser Ordner enthält alle Ergebnisdateien des Gastaufenthalts, inklusive Grafiken.	**Skripte und Ergebnisdateien** werden **nicht** automatisch zwischen **JoSuA** und dem **Projektverzeichnis** abgeglichen.
doc	Dieser Ordner dient zur Erstellung von Notizen während des Gastaufenthalts. Diese Dokumente stehen nur im Gastaufenthalt zur Verfügung und können nicht herausgegeben werden.	

Wenn Sie in Ihrem Projekt die Datenfernverarbeitung per Remote-Desktop nutzen, dann enthält Ihr Projektverzeichnis weitere Unterverzeichnisse. Diese werden in Kapitel 6 beschrieben.

2.2 Weitere Ressourcen

2.2.1 Ado-Files (Stata)

- Ado-files können nicht aus dem Internet heruntergeladen werden. Das FDZ stellt die Pakete des Statistical Software Components (SSC) Archive (im Original gehostet am Boston College Department of Economics (BOCODE) und bereitgestellt von RePEc) in seinem Netzwerk zur Verfügung (siehe Liste unter https://doku.iab.de/fdz/access/stata_ado.pdf).
- Die Pakete können in der FDZ-Infrastruktur genutzt werden. Verwenden Sie zur Installation der Pakete folgenden Befehl, der alle zugehörigen Dateien in Ihr **prog**-Verzeichnis kopiert:

```
fdzinstall packagename
```

Dieser Befehl kopiert alle relevanten Dateien in das **prog**-Verzeichnis des jeweiligen Projekts. Der Befehl sollte pro Projektverzeichnis nur einmal verwendet werden, da wiederholte Ausführungen zu unfreiwilligen Updates der ados führen können. Die Pakete stehen anschließend im **prog**-Verzeichnis bereit und sind automatisch in Stata eingebunden. Die Installation und Verwendung der Ado-Files im Ordner prog stehen im Gastaufenthalt und via JoSuA gleichermaßen zur Verfügung.

- Die im FDZ-Gästenetz bereitgestellte Ado-File-Sammlung wird regelmäßig aktualisiert. Dabei werden die von Nutzenden in **prog**-Verzeichnisse kopierten Dateien vom FDZ nicht aktualisiert. Um eine etwaige neuere Version eines ado-Pakets ins prog-Verzeichnis zu kopieren, muss der Befehl

```
fdzinstall packagename
```

erneut ausgeführt werden. Die vorherige Version des Ado-Files wird dann ersetzt.

Um bei etwaigen Problemen durch Updates auf die vorherige Version zurückzukehren, bietet fdzinstall die repo()-Option, mit welcher ein vorheriger Stand der Ado-File-Sammlung aufgerufen werden kann. Eine Liste verfügbaren Versionen findet sich hier (https://doku.iab.de/fdz/access/stata_ado.pdf).

- Wenn Sie Ado-Files verwenden möchten, die nicht Teil des SSC-Archives sind bzw. zu denen eine neue Version noch nicht am FDZ zur Verfügung steht, laden Sie diese bitte in JoSuA hoch (nur Textdateien wie bspw. .ado, .do, …). Ado-Files werden im Reiter *Projects* (siehe Kapitel 10.4), unter *Resources*, einmalig hochgeladen. Sie stehen dann für die Datenfernverarbeitung in JoSuA zur Verfügung. Hier hochgeladene Ado-Files können danach auch in die Gastaufenthalts- und Remote-Desktop-Umgebungen heruntergeladen und verwendet werden.
- Wenn Sie Ado-Files verwenden möchten, die nicht per Resources hochgeladen werden können (z.B. weil sie .mlib oder .plugin Dateien enthalten), schicken Sie diese mindestens drei Arbeitstage vor Beginn des Gastaufenthalts per E-Mail an das FDZ-Postfach. Auch diese Ado-Files werden im Ordner **prog** gespeichert.
- .mo-Dateien können leider nicht geprüft werden. Bitte erzeugen Sie derartige Dateien im Gästenetz des FDZ neu. Bei Ado-Files gibt es dazu i.d.R. Do-Files, die die Objekte definieren und `"function"` oder `"fun"` im Namen tragen. Folgendermaßen können .mo-Dateien via JoSuA (unabhängig vom Modus) in Ihrem Projektverzeichnis erzeugt werden:
  - Erstellen/Anpassen des Do-Files, das die .mo-Datei erzeugt; dabei ist zu beachten:
    - Speichern des erstellten Objekts mit einem Befehl wie diesem:
      `mata mosave xyz(),replace dir($localprog/x)‘`
    - Dabei steht `xyz` für den Namen des erstellten Objekts. Der Buchstabe hinter `"$localprog/"` steht für einen etwaigen Unterordner in dem **prog**-Ordner in Ihrem Projektverzeichnis, in dem das Ado-File die .mo-Datei ggf. erwartet (also z.B. `"$localprog/x"` für eine Datei `xyz.mo`).
    - Erzeugen Sie den ggf. nötigen Unterordner in Ihrem **prog**-Ordner vor dem Speichern der .mo-Datei mit einem solchen Befehl:
      `capture mkdir $localprog/x`
  - Ausführen des relevanten Do-files z.B. durch Aufruf in einem master.do via JoSuA

– Falls Sie die .mo-Dateien während eines Gastaufenthalts erstellen wollen, nutzen Sie bitte `$prog/x` statt `$localprog/x`.

2.2.2 R-Packages

R-Packages können nicht aus dem Internet heruntergeladen oder über JoSuA hochgeladen werden. Bitte wenden Sie sich per E-Mail an das FDZ, wenn Sie R-Packages benötigen, die nicht im Gästenetz verfügbar sind.

2.2.3 Zuspielen externer aggregierter Datensätze

- Externe Daten ohne jeden Personen- oder Betriebsbezug (z.B. Anteil landwirtschaftlicher Fläche an Gesamtfläche eines Kreises, Distanz zwischen Hauptstadt eines Herkunftslandes und Berlin, tagesgenauer Mondkalender) sind unbedenklich und können immer zugespielt werden.
- Das Zusammenspielen von personen-/betriebsbezogenen Einzelbeobachtungen ist nicht möglich bzw. nicht erlaubt!
- Externe Kenngrößen auf aggregiertem Niveau (z. B. Arbeitslosenquoten nach Kreisen) dürfen den Daten zugespielt werden, wenn sie den Datenschutzrichtlinien des FDZ entsprechen. Jeder aggregierte Wert muss auf mindestens 3 Einzelbeobachtungen beruhen (im Beispiel also mindestens 3 Arbeitslose und 3 Nicht-Arbeitslose). Bei Aggregatgrößen aus Quellen der amtlichen Statistik ist dies in der Regel der Fall. Der Datensatz muss Variablen enthalten, die die Anzahl der zugrundeliegenden Beobachtungen angeben. Bitte geben Sie ausschließlich absolute Häufigkeiten an (z.B. Anzahl Männer und Anzahl Frauen anstatt des Frauenanteils). Relative Häufigkeiten lassen eine Überprüfung der Einhaltung des Datenschutzes nicht ohne Weiteres zu. Die Berechnung von Anteilswerten sollte erst in der FDZ-Umgebung erfolgen.
- Die Kenngrößen **müssen** dem FDZ mit einer Beschreibung des Datensatzes (inkl. der Merkmale, der Aggregationsebene und der Quellenangabe) als **Stata-Datensatz** übermittelt werden (iab.fdz@iab.de). Bitte komprimieren Sie den aggregierten Datensatz vor der Übermittlung an das FDZ mit dem Befehl `compress`. Eingeschickte Datensätze im Excel-Format können nicht zur Verfügung gestellt werden. Die Übermittlung aggregierter Kennzahlen innerhalb von Skripten ist nicht zulässig.
- Nach Prüfung und Freigabe durch das FDZ werden die externen Datensätze im Verzeichnis **orig** zur Verfügung gestellt. Bitte stimmen Sie die Zuspielung von Aggregatdaten frühzeitig mit dem FDZ ab. Die Bereitstellung externer aggregierter Daten in Ihrem Projektverzeichnis kann nach erfolgreicher Prüfung bis zu **3 Arbeitstage** dauern.

- Sollten Sie Datensätze aus den zentralen Arbeitshilfen des FDZ benötigen, dann schreiben Sie bitte eine E-Mail mit den Bezeichnungen der gewünschten Datensätze an das FDZ-Postfach. Wir stellen Ihnen die Datensätze dann innerhalb von drei Arbeitstagen in Ihrem **orig** Ordner zur Verfügung. Die verfügbaren Arbeitshilfen des FDZ sind auf unserer Website unter „Zentrale Arbeitshilfen" einzusehen.

2.3 Einrichtung einer Testumgebung am eigenen Arbeitsplatz

Zur Entwicklung und zum Testen von Skripten empfehlen wir, eine Testumgebung am eigenen Arbeitsplatz aufzubauen, um die Datenfernverarbeitung mit JoSuA oder Ihren Gastaufenthalt bestmöglich vorzubereiten. Im Folgenden erfahren Sie, wie Sie sich außerhalb der FDZ-Infrastruktur eine Testumgebung einrichten, die dem Projektverzeichnis auf den FDZ-Servern entspricht. Wenn Sie auch den Remote-Desktop-Zugang nutzen, können Sie diesen Abschnitt evtl. überspringen.

2.3.1 Einrichtung eines Projektverzeichnisses

- Richten Sie auf Ihrem PC ein Verzeichnis „**fdz[Ihre Projektnummer]**" ein.
- Erstellen Sie in diesem Verzeichnis die Ordner **orig**, **prog**, **data** und **log**.
- Erstellen Sie keine Unterordner.
- In JoSuA und im Gastaufenthalt sind die Globals für Pfadangaben automatisch definiert. Für Ihre eigene Testumgebung müssen Sie die Globals **$orig**, **$data**, **$prog**, **$log** und ggf. den **adopath** definieren, bevor Sie die Datei **master.do** starten. Schreiben Sie diese Global-Definitionen **nicht** in master.do, sondern in eine Datei `profile.do` (siehe FDZ-Templates). Diese wird von Stata automatisch ausgeführt und muss nicht durch master.do angesprochen werden. Sie kann beispielsweise im aktuellen Arbeitsverzeichnis abgelegt werden (siehe auch http://www.stata.com/help.cgi?profile). Die Datei **profile.do** darf **nicht in JoSuA** hochgeladen werden.
- Sollten Sie für Ihr Projekt mehr als ein Datenprodukt nutzen, erstellen Sie für jeden Datensatz wie oben beschrieben ein eigenes Projektverzeichnis. Bitte beachten Sie, dass am FDZ kein Datentransfer (d.h. Kopieren oder Verschieben) zwischen diesen Projektverzeichnissen durch die Nutzenden möglich ist.

2.3.2 Download und Vorbereitung der Testdaten

- Für die meisten durch das FDZ angebotenen Datensätze stellen wir Testdaten zur Verfügung. Diese haben die gleiche Datenstruktur wie die Originaldaten, können aber

nicht für Analysen verwendet werden. Bitte beachten Sie die Hinweise zu den jeweiligen Testdaten auf unserer Website.

- Speichern Sie die Testdaten im Ordner **orig**.
- Viele FDZ-Datenprodukte enthalten sensible Merkmale, die gesondert beantragt werden müssen. In den Testdaten sind grundsätzlich alle sensiblen Merkmale enthalten. Sie müssen die Testdaten daher vorab anpassen, je nachdem welche sensiblen Merkmale Sie beantragt haben (siehe FDZ-Template `prepare_test_environment.do`).
- Für Erweiterungsmodule werden keine Testdaten zur Verfügung gestellt. Sie können diese ergänzenden Datensätze anhand der Datensatzbeschreibung im Datenreport selbst erstellen. Hierzu müssen Sie beispielsweise Variablen mit Zufallszahlen füllen (siehe FDZ-Template prepare_test_environment.do).

2.3.3 Verwendung der FDZ-Templates

- Die FDZ-Templates können von der FDZ-Website heruntergeladen werden. Diese sind ausschließlich im Stata-Format verfügbar. Sollten einzelne Berechnungen nicht mit Stata durchführbar sein, halten Sie sich bitte dennoch im Rahmen der anderen Software (z.B. R) an die Grundstruktur dieser Templates.
- Speichern Sie die Skripte im Ordner **prog**.
- Die FDZ-Templates wurden mit dem SIAB-Testdatensatz entwickelt. Sie können die Einrichtung der Testumgebung und die FDZ-Templates testen, indem Sie die SIAB-Testdaten von unserer Website herunterladen. Für andere Datensätze müssen Sie die einzelnen Skripte selbst anpassen.
- Weitere Informationen zur Nutzung der FDZ-Templates zur Vorbereitung des Arbeitens mit den Testdaten am eigenen Arbeitsplatz und als Basis für die Gestaltung eigener Skripte finden Sie im Zip-Archiv der FDZ-Templates.

2.3.4 Programmpakete und externe Aggregatdaten

- Um die Zuspielung und Analyse externer Aggregatdaten zu testen, speichern Sie diese Dateien in Ihrer Testumgebung im Ordner **orig**.
- Wenn Sie Programmpakete nutzen möchten, müssen diese im Ordner **prog** gespeichert werden. Siehe hierzu Kapitel 2.2.1.[1]

[1] Stata sucht standardmäßig in verschiedenen Verzeichnissen nach Ado-Files. Um zu vermeiden, dass beim Testen der Programme Ado-Files eingebunden sind, die im Gastaufenthalt oder bei der Datenfernverarbeitung nicht zur Verfügung stehen, sollte der Suchpfad für die Testumgebung auf **$prog** eingeschränkt werden (siehe FDZ-Template `profile.do`).

2.3.5 Testen der Skripte vor Gastaufenthalt und Datenfernverarbeitung mit JoSuA

- Überprüfen Sie Ihre Auswertungsprogramme vor Gastaufenthalt und Datenfernverarbeitung mit JoSuA mit Hilfe der auf der FDZ-Website zur Verfügung gestellten Testdaten.
- Führen Sie die Datei `master.do` in Ihrer Testumgebung aus.
- Bitte beachten Sie, dass die Dateigrößen der Testdaten nur einem Teil der Originaldaten entsprechen. Die Originaldaten und aufbereiteten Datensätze sind deutlich größer und Programme laufen dementsprechend länger.
- Nach erfolgreichem Test laden Sie Ihre Programme **unverändert** in JoSuA im Internal Use Modus hoch.

3 Gestaltung der Skripte

Die Gestaltung der Skripte bzw. Programmdateien ist zentral für eine reibungslose und schnelle Abarbeitung der Datenschutzprüfungen durch das FDZ-Team. Daher ist es essentiell, dass sich alle Nutzenden an diese Richtlinien halten.

Sie finden ausführliche Erläuterungen in der Arbeitshilfe „Richtlinien“.

Des Weiteren sollten Nutzende die **FDZ-Templates** als Vorlagen für Ihre Programmdateien verwenden. Sie finden die FDZ-Templates auf unserer Website.

Bitte beachten Sie, dass Skripte keine Daten enthalten dürfen. Kopieren Sie daher niemals Ergebnisse, Betriebs- oder Personennummern in Ihre Skripte. Externe Daten müssen als Stata-Datensatz ans FDZ geschickt werden (siehe Kapitel 2.2.3).

Bitte verwenden Sie bei der Benennung der Skript-Dateien nur folgende Zeichen: 0-9, a-z, A-Z, -,_.

4 Datennutzung mit JoSuA (Job Submission Application)

Für die Nutzung der meisten FDZ-Datenprodukte ist die kontrollierte Datenfernverarbeitung essentiell. Verträge für Gastaufenthalte oder den Remote-Desktop-Zugang schließen immer auch die Nutzung der (indirekten) kontrollierten Datenfernverarbeitung mit ein. Welcher Datenzugang für Ihr Projekt vorliegt, entnehmen Sie bitte Ihrem Nutzungsvertrag. Für alle Zugangswege setzt das FDZ die webbasierte Software **JoSuA** (Job Submission Application) ein:

- Im Rahmen der **kontrollierten Datenfernverarbeitung** dient JoSuA zum Hochladen der auszuführenden Skripte, zur Einsichtnahme in vorläufige Ergebnisse und zum Herunterladen der geprüften und freigegebenen Resultate.
- Im Rahmen von **Gastaufenthalten** müssen Skripte, die vorab am eigenen Arbeitsplatz erstellt wurden, in JoSuA (via „Transfer") hochgeladen werden. Nach erfolgter Prüfung lassen sich diese vom Gästenetz des FDZ aus herunterladen und dort weiterbearbeiten. Zudem können Skripte, die während eines Gastaufenthalts erstellt bzw. verändert wurden, über JoSuA hochgeladen werden, damit sie anschließend für die Datenfernverarbeitung und die Weiterverarbeitung am eigenen Arbeitsplatz zur Verfügung stehen. Die Ergebnisse Ihres Gastaufenthalts erhalten Sie ausschließlich über kontrollierte Datenfernverarbeitung. Weitere Informationen zum Gastaufenthalt sowie zur Nutzung von JoSuA im Kontext von Gastaufenthalten finden Sie in Kapitel 5.
- Im Rahmen der Datenfernverarbeitung per **Remote-Desktop** dient JoSuA, wie bei Gastaufenthalten, zum Transfer von Skripten. Außerdem ist JoSuA nötig, um die im direkten Remote-Zugriff erstellten vorläufigen Ergebnisse zu validieren (siehe Kapitel 6.3).

4.1 Grundsätzliches zur Nutzung von JoSuA

- Das FDZ nutzt JoSuA für:
  - die kontrollierte Datenfernverarbeitung (siehe Kapitel 10.5),

  - den Transfer von Skripten[1] vor und nach Gastaufenthalten (siehe Kapitel 10.7) und
  - das Hochladen von Ado-Files (siehe Kapitel 2.2.1 und Kapitel 10.4).

- Zugangsdaten:
  - Sie erhalten Ihre Kennung für JoSuA sowie den ersten Teil Ihres Passwortes per E-Mail.
  - Der zweite Teil des Passwortes wird Ihnen per Brief zugeschickt.
  - Bitte beachten Sie, dass Sie für den Zugang im Gastaufenthalt bzw. Remote-Desktop eigene Kennungen bekommen (siehe Kapitel 5.2 und Kapitel 6.1). Das dazugehörige Initialpasswort ist identisch mit dem für JoSuA.[2]

- Kontrollierte Datenfernverarbeitung mit JoSuA:
  - Hierbei kann auf Daten in den Verzeichnissen **orig** und **data** zurückgegriffen werden. Beachten Sie hierzu die Programmierbeispiele in den FDZ-Templates.
  - Skripte und Ergebnisdateien (z.B. Log-Files, Grafiken) werden von JoSuA nicht in die Verzeichnisse **prog** und **log** sondern in eine separate Ablage gespeichert. Die Dateien, die in JoSuA im Ordner **data** gespeichert werden, werden auch in das Verzeichnis **data** geschrieben. Umgekehrt sind die während des Gastaufenthalts/Remote-Desktops in den Verzeichnissen **prog** und **log** gespeicherten Dateien nicht während der Nutzung von JoSuA verfügbar.
  - Alle **Ado-Files** im Ordner **prog** Ihres Projektverzeichnisses am FDZ sind bereits automatisch eingebunden.

- Externe aggregierte Daten können nicht über JoSuA hochgeladen werden, sondern müssen per E-Mail an das FDZ geschickt werden. Ausführliche Informationen zum Vorgehen finden Sie in Kapitel 2.2.3.

Die Bedienung von JoSuA wird detailliert in Kapitel 10 beschrieben.

Aktuelle Antworten auf häufige Fragen zu JoSuA finden Sie auch auf der JoSuA-Seite der FDZ-Website. Dort finden Sie auch ein Video-Tutorial, das Ihnen den Einstieg zur Nutzung von JoSuA erleichtern soll.

[1] Obwohl im Rahmen der kontrollierten Datenfernverarbeitung über JoSuA derzeit nur Stata genutzt werden kann, kann JoSuA dennoch zum Transfer von Skripten anderer Statistik-Pakete (wie z.B. R oder Octave) genutzt werden.

[2] Änderungen des Passwortes in JoSuA werden nicht mit der Gastaufenthalts- bzw. Remote-Desktop-Kennung synchronisiert.

4.2 Internal Use Modus und Presentation / Publication Modus

Grundsätzlich stehen in JoSuA zwei Nutzungsmodi zur Verfügung:

1) Internal Use

- Der Internal Use Modus dient zur Datenaufbereitung, Vorbereitung von Gastaufenthalten und zum Testen der Skripte. Bitte verwenden Sie den Internal Use Modus solange für Ihre Auswertungen, bis Sie Ihre Endergebnisse oder Zwischenergebnisse für Präsentationen oder Publikationen erstellt haben.
- Überprüfen Sie vorher die Lauffähigkeit Ihrer Programme in Ihrer Testumgebung.
- Die Ergebnisse sind nur für den „internen" Gebrauch, d. h. sie dienen nur zur Entwicklung der Skripte.
- Hochgeladene Aufträge werden automatisch gestartet und die Ergebnisse können in der JoSuA-Webschnittstelle betrachtet, aber nicht heruntergeladen werden.
- Die für Sie sichtbaren Ergebnisse werden einer automatischen skriptbasierten Datenschutzprüfung unterzogen und ggf. zensiert.
- Der Internal Use Modus dient ausschließlich der Einsichtnahme zur Weiterentwicklung Ihrer Skripte. Jede andere Verwendung der Ergebnisse ist untersagt. Die Ergebnisse dürfen nicht abgeschrieben, fotografiert, ausgedruckt, kopiert oder anderweitig entnommen werden. Es ist außerdem untersagt, Dritten, die nicht im Nutzungsvertrag des Projekts genannt sind, Einsicht zu gewähren. Übermitteln Sie bitte **NIEMALS** Ergebnisse, die im Internal Use Modus erstellt wurden, oder sonstigen nicht vom FDZ-Team freigegebenen Output! Die Übermittlung von nicht freigegebenen Ergebnissen – auch an das FDZ – stellt eine Vertragsverletzung dar. Sollten Sie Fragen zu einem Internal Use Job haben, dann senden Sie uns bitte ausschließlich die Projekt- und Jobnummer zu.
- Bitte beachten Sie, dass zur Gewährleistung des Datenschutzes im Internal Use Modus einige Befehle grundsätzlich gesperrt werden[3]. Ausnahmeregelungen sind nicht möglich. Sie haben in diesen Fällen die Möglichkeit, durch Auswahl des Presentation / Publication Modus manuell geprüfte Ergebnisse zu erhalten.
- Um eine Rückrechnung und Ergänzung von Ergebnissen auszuschließen, sind die Ergebnisse vorheriger Aufträge im Internal Use Modus nicht mehr verfügbar, sobald innerhalb des gleichen Projekts ein Auftrag im Presentation / Publication Modus eingeschickt wird.

2) Presentation / Publication

[3] Derzeit ist der Befehl list im Internal Use Modus gesperrt. Das FDZ behält sich die Sperrung weiterer Befehle vor.

- In diesem Modus werden herunterladbare Log-Dateien und Grafiken erstellt.
- Der Modus ist auszuwählen, wenn Ergebnisse im Rahmen einer Präsentation oder Publikation veröffentlicht werden sollen.
- Überprüfen Sie vorher die Lauffähigkeit Ihrer Programme und die Einhaltung unserer Richtlinien im Internal Use Modus.
- Beschränken Sie sich auf die Ergebnisse, die Sie für eine Veröffentlichung benötigen. Laden Sie nur die Skripte hoch, die die für die jeweilige Präsentation oder Publikation notwendigen Ergebnisse erzeugen. Kommentieren Sie Befehle aus, deren Ergebnisse nicht benötigt werden. Es kann auf alle vorher im Internal Use Modus erstellten Datensätze zugegriffen werden, d. h. Aufbereitungsschritte müssen nicht erneut laufen.
- Stellen Sie bitte sicher, dass keine Befehle wie `nostop` in Ihren Programmen enthalten sind, die eigentlich auftretende Programmabbrüche umgehen.
- Stellen Sie sicher, dass Ihre Ergebnisdarstellung mit unseren Richtlinien konform ist. In den Richtlinien finden Sie Hinweise zur Ausgabe von Tabellen, Grafiken und aggregierten Datensätzen.
- Orientieren Sie sich an den FDZ-Templates zur effizienten Programmierung und Ergebnisdarstellung. So stellen Sie sicher, dass Ihre Programme möglichst schnell abgeschlossen werden und die Datenschutzprüfung durch das FDZ zügig erfolgen kann.

Eine detaillierte Anleitung zur Nutzung von JoSuA und Hinweise zu Job-Abbrüchen in JoSuA finden Sie in Kapitel 10.

4.3 Ablauf der kontrollierten Datenfernverarbeitung mit JoSuA

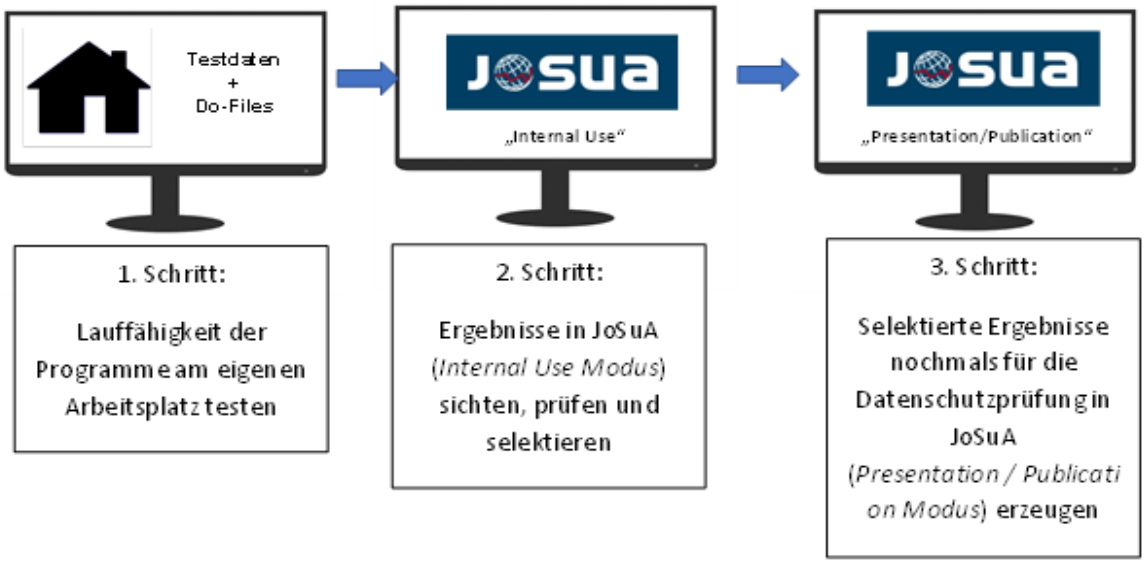
Einen allgemeinen Überblick über den Arbeitsablauf finden Sie in Abbildung 1.

- Die von Ihnen in JoSuA hochgeladenen Skripte müssen **vollständig lauffähig** sein. Das FDZ führt keine Fehlerbehebungen durch. Testen Sie Ihre Programme daher immer vorher in Ihrer Testumgebung (siehe Kapitel 2.3) oder (falls verfügbar) im Remote-Desktop.
- Stellen Sie sicher, dass in Ihren Skripten die Richtlinien umgesetzt sind.
- Sollten Sie **Ado-Files** in JoSuA benötigen, beachten Sie bitte die Hinweise in Kapitel 2.2.1 unserer Anleitungen. Alle Ado-Files im Ordner **prog** Ihres Projektverzeichnisses am FDZ sind bereits automatisch eingebunden.

- Beachten Sie die technischen Restriktionen bezüglich Anzahl und Größe von Programm- und Ergebnisdateien sowie die Speicherbegrenzungen (siehe Kapitel 10.8). Grundsätzlich bitten wir Sie, nicht unnötig Arbeitsspeicher zu belegen, da bei einer hohen Auslastung der Rechenserver die Gefahr besteht, dass Ihr Job und die Jobs anderer Nutzender abbrechen oder stark ausgebremst werden.
- Laden Sie Ihre Programme immer zuerst im Internal Use Modus hoch, um die Ergebnisse zu sichten, prüfen und zu selektieren. Stellen Sie zudem sicher, dass die Ergebnisse, die Sie später nochmals im Presentation / Publication Modus erzeugen möchten, den datenschutzrechtlichen Anforderungen entsprechen (siehe Richtlinien).
- Ihre Auswertungsprogramme werden mit der aktuellsten am FDZ vorhandenen Stata-Version gestartet.
- Bitte warten Sie mit dem Hochladen der nächsten Auswertung, bis die Ergebnisse des letzten vorangegangenen Jobs innerhalb Ihres Projekts bereitgestellt wurden. Das FDZ behält sich vor, gleichzeitig laufende Jobs ohne Rücksprache mit den Nutzenden abzubrechen, falls durch sie die Jobs anderer Nutzender beeinträchtigt werden. Wie viele Jobs pro Tag/Woche im Internal Use Modus sowie Presentation / Publication Modus hochgeladen werden können, kann der FDZ-Website entnommen werden. Auf dieses Limit werden auch Jobs angerechnet, die direkt aufgrund fehlerhafter Programmierungen abbrechen.
- Der Zeitpunkt der Freigabe des Outputs ist abhängig von der Serverauslastung sowie dem Rechen- und Prüfaufwand.

Abbildung 1 fasst die zentralen Schritte der kontrollierten Datenfernverarbeitung mit JoSuA nochmals kurz zusammen.

Abbildung 1: Ablaufplan kontrollierte Datenfernverarbeitung mit JoSuA



5 Gastaufenthalte

Das FDZ empfiehlt für alle Projekte, die komplexe Daten wie z.B. SIAB, Linked-Employer-Employee-Daten des IAB (LIAB) oder andere verknüpfte Datenprodukte analysieren, mindestens vor der ersten Nutzung von JoSuA einen Gastaufenthalt durchzuführen (oder alternativ den Remote-Desktop zu nutzen, siehe Kapitel 6). Dabei kann man sich schneller mit den Daten vertraut machen und die teilweise sehr aufwändige Aufbereitung von Analysedatensätzen direkt vornehmen.

5.1 Vorbereitungen für den Gastaufenthalt

- Nach Abschluss eines Nutzungsvertrages erhalten Sie Ihre Kennung per E-Mail und das Initialpasswort gesplittet per Brief und Mail. Sollten später noch weitere Projekte dazukommen, ist das Initialpasswort immer das gleiche wie beim ersten Projekt. Bitte beachten Sie, dass Sie für die Nutzung von JoSuA eine eigene Kennung bekommen (siehe Kapitel 4.1). Das dazugehörige Passwort ist identisch mit dem Initialpasswort für Gastaufenthalte.
- Vereinbaren Sie im Voraus einen **Termin** zum Gastaufenthalt im FDZ oder einem der externen Standorte. Berücksichtigen Sie, dass die Kapazitäten an Gastarbeitsplätzen begrenzt sind, nicht jeder Terminwunsch erfüllbar ist und die Buchungsmodalitäten je nach Standort unterschiedlich geregelt sind. Informationen dazu finden Sie auf unserer Website (Kontaktdaten).
- Wenn Sie während Ihres Gastaufenthalts am FDZ vorbereitete **Auswertungsprogramme** nutzen wollen, laden Sie diese **spätestens 3 Arbeitstage** vor dem ersten Aufenthaltstag über JoSuA über die Funktion „Transfers" hoch.
- Alle hochgeladenen Skripte können Sie nach erfolgter Eingangsprüfung durch das FDZ während des Gastaufenthalts von einem internen Server herunterladen und im Ordner **prog** Ihres Projektverzeichnisses abspeichern.
- Bitte beschränken Sie sich bei dem Upload auf Auswertungsprogramme, die **neu** sind oder seit Ihrem letzten Gastaufenthalt **verändert** wurden. Der Transfer von unveränderten Programmen erzeugt unnötigen Prüfaufwand beim FDZ-Personal. Skripte, die das FDZ zur Datenaufbereitung zur Verfügung stellt, sind im Gästenetz abgelegt und müssen nicht über Transfers hochgeladen werden. Dazu gehören alle

Programmdateien aus FDZ-Methodenreporten, die FDZ-Stata-Templates, das Panel-Tool für das IAB-Betriebspanel, die Do-Files von Dauth et al. (2023) und weitere Arbeitshilfen.

- **Aggregatdatensätze**, die Sie für Ihren Gastaufenthalt benötigen, müssen ebenfalls **spätestens 3 Arbeitstage** vor dem ersten Aufenthaltstag zur Prüfung an das FDZ gesendet werden. Nennen Sie bitte das Datum des geplanten Gastaufenthalts in der E-Mail. Falls Sie innerhalb von zwei Arbeitstagen keine Eingangsbestätigung erhalten haben, geben Sie uns bitte Bescheid. Beachten Sie den Abschnitt zu externen Aggregatdaten (Kapitel 2.2.3).
- Wollen Sie im Gastaufenthalt **Ado-Files** nutzen, beachten Sie bitte Kapitel 2.2.1.

5.2 Während des Gastaufenthalts

- Zur Erstanmeldung im Gastaufenthalt verwenden Sie bitte Ihr Initialpasswort, dass Sie gesplittet per Brief und Mail erhalten haben (siehe Kapitel 5.1).
- Bitte ändern Sie Ihr Passwort nach der Erstanmeldung. Die Anleitung zur Passwortänderung finden Sie bei den Working Tools (unter **Vorgaben_Guidelines** „**Shortguide Password**").
- **FDZ-Arbeitshilfen**:
  - Sie haben im Gästenetz Zugriff auf ein allgemeines Verzeichnis („**Working_Tools**"). Dort finden Sie:
    - FDZ-Datenreporte
    - weitere Arbeitshilfen und zentrale Methodenreporte inkl. dazugehöriger Dateien (Datensätze und Skripte)
    - diese Anleitungen, die Richtlinien und die FDZ-Templates
    - das Ado-Repositorium
  - Sie können sich die dort verfügbaren Dateien im Gastaufenthalt in das eigene Verzeichnis kopieren (in die Ordner prog, data und doc).
  - Die verfügbaren Arbeitshilfen des FDZ sind auch auf unserer Website unter „Zentrale Arbeitshilfen" einzusehen.
- Auf Datensätze, die im Rahmen der kontrollierten Datenfernverarbeitung im Ordner **data** erstellt wurden, können Sie im Gastaufenthalt zugreifen. Datensätze, die ausschließlich im Remote-Desktop-Zugang erzeugt werden, müssen im Gastaufenthalt nochmals generiert werden.

5.3 Erhalt der geprüften Ergebnisse nach Gastaufenthalt

Am Ende des Gastaufenthalts

- Bitte laden Sie Ihre Skripte zum Ende Ihres Gastaufenthalts in JoSuA über die Funktion „Transfers" hoch. So stehen Ihnen die Programme für die weitere Bearbeitung am eigenen Arbeitsplatz oder für die Datenfernverarbeitung zur Verfügung. Sie finden im Gastaufenthalt einen Link zu JoSuA auf dem Desktop des Stata-Servers.
- Bitte beschränken Sie sich beim Transfer der Skripte auf solche, die neu sind oder während des Gastaufenthalts verändert wurden. Der Transfer von unveränderten Programmen erzeugt unnötigen Prüfaufwand beim FDZ-Personal. Die im Verlauf des Gastaufenthalts erzeugten Ergebnisse werden nicht herausgegeben. Falls Sie die Ergebnisse sofort für Publikationen oder Präsentationen benötigen, laden Sie alle notwendigen Skripte **zusätzlich** in JoSuA im Presentation / Publication Modus hoch. Sorgen Sie dafür, dass master.do und alle Programmdateien für den Presentation / Publication Modus vollständig lauffähig sind.
- Das bedeutet, wenn Sie sowohl Ihre Skripte als auch Ergebnisse nach dem Gastaufenthalt erhalten möchten, dass Sie Ihre Skripte einmal unter „Transfers" und einmal unter „Submission" im Presentation / Publication Modus hochladen müssen. Unter „Transfers" können Sie allerdings zusätzlich auch neue oder geänderte Skripte hochladen, die nicht Teil des im Presentation / Publication Modus gestarteten Jobs sind.
- Alle hochgeladenen Skripte und ggf. Ihre Ergebnisse stehen nach erfolgter Datenschutzprüfung in JoSuA zum Download zur Verfügung. Eine nachträgliche Übermittlung von Skripten durch das FDZ ist nicht möglich.
- Bitte melden Sie sich am Ende Ihres Aufenthaltes korrekt vom Server ab (s. Handbuch Gastaufenthalt im Gästenetz).

6 Datenfernverarbeitung per Remote-Desktop

Wenn Sie für Ihren Datenzugang den Remote-Desktop gewählt haben, kann dieser alternativ oder zusätzlich zum Gastaufenthalt genutzt werden, um Ihre Analysen für JoSuA vorzubereiten. Die Nutzung des Remote-Desktops vom eigenen Büro aus ist dabei in vielen Bereichen analog zum Zugang im Gastaufenthalt, daher beschränkt sich dieser Abschnitt auf Unterschiede im Vergleich zu Kapitel 5 und Punkte, die zusätzlich zu beachten sind.

6.1 Vorbereitungen für den Remote-Desktop

- Eine **technische Bedienungsanleitung** für den Remote-Desktop finden Sie in Kapitel 11.
- Bitte beachten Sie, dass Sie für die Nutzung des Remote-Desktop eine eigene Kennung bekommen, die sich von der für JoSuA und den Gastaufenthalt unterscheidet (siehe Kapitel 4.1 und Kapitel 5.1). Das dazugehörige Passwort ist identisch mit dem Initialpasswort für Gastaufenthalte.
- Zur Erstanmeldung im Remote-Desktop verwenden Sie bitte Ihr Initialpasswort, dass Sie gesplittet per Brief und Mail erhalten haben.
- Bitte ändern Sie Ihr Passwort nach der Erstanmeldung. Die Anleitung zur Passwortänderung finden Sie als Shortcut auf dem Desktop und bei den Working Tools (unter **Vorgaben_Guidelines** „**Shortguide Password**").
- Steht der Remote-Desktop zu Verfügung, sind in Ihrem Projektverzeichnis nicht nur die Ordner **prog**, **orig**, …, sondern jeweils auch die Zwillingsordner **prog_remote**, **orig_remote**, …, vorhanden.
- Während Sie den Remote-Desktop nutzen, haben Sie nur Zugriff auf die **_remote** Ordner, bei der Nutzung von JoSuA oder im Gastaufenthalt nur auf die Ordner ohne **_remote** in der Bezeichnung.
- Bei der Programmierung Ihrer Skripte macht das keinen Unterschied. Benutzen Sie bitte weiterhin die in den Templates und Anleitungen genannten Globals. Die Zuordnung erfolgt automatisch.
- Die Zwillingsordner werden von uns nicht synchronisiert.

- Wollen Sie Skripte, die Sie außerhalb des FDZ-Gästenetzes entwickelt haben, im Remote-Desktop nutzen, so müssen Sie diese in JoSuA (*„Transfers"*) hochladen, und nach erfolgter Eingangsprüfung durch das FDZ in der Remote-Desktop-Umgebung wieder herunterladen und im Ordner **prog_remote** Ihres Projektverzeichnisses abspeichern.
- Die Bereitstellung von Ado-Files erfolgt entsprechend Kapitel 2.2.1.
  - Die Bereitstellung externer Datensätze erfolgt standardmäßig parallel in **orig** und **orig_remote**.

6.2 Arbeiten mit dem Remote-Desktop

- Aus Datenschutzgründen werden in den bereitgestellten FDZ-Datenprodukten sensible Merkmale zum Arbeiten im Remote-Desktop verfremdet. Eine Liste der betroffenen Merkmale können Sie den jeweiligen Datenreporten entnehmen.
- Die Verfremdung findet lediglich auf der am stärksten disaggregierten Ebene statt. Beispielsweise wird der Wohnort-Kreis verfremdet, während das Wohnort-Bundesland unverändert bleibt.
- Im Remote-Desktop erzeugte Ergebnisse, die auf sensiblen Merkmalen aufbauen, sind daher als vorläufig zu betrachten. Wir empfehlen, wichtige Programmschritte und Ergebnisse regelmäßig im Internal Use Modus von JoSuA zu validieren.
- Wenn Sie keine sensiblen Merkmale beantragt haben, sind **orig** und **orig_remote** identisch.
- Datenschutzregeln gelten analog zum Gastaufenthalt: Daten und Ergebnisse dürfen nicht abgeschrieben, fotografiert, ausgedruckt, kopiert oder anderweitig entnommen werden. Es ist außerdem untersagt, Dritten, die nicht im Nutzungsvertrag des Projekts genannt sind, Einsicht zu gewähren. Jede Verwendung dieser Ergebnisse außer zur Weiterentwicklung Ihrer Programme stellt eine Vertragsverletzung dar.
- Bitte melden Sie sich korrekt vom Gästenetz-Server des FDZ ab, wenn Ihre Arbeit mit dem Remote-Desktop beendet ist, und vermeiden Sie offene Sessions, an denen Sie länger nicht arbeiten. Andere Forschende werden Ihnen den Performancegewinn danken!

6.3 Validierung und Export von Ergebnissen

- Im Remote-Desktop erzeugte Ergebnisse werden nicht herausgegeben und müssen evtl. validiert werden.

- Nutzen Sie zur Validierung den Internal Use Modus von JoSuA. Laden Sie hierzu alle relevanten Skripte vom Remote-Desktop aus unter „Submission" hoch und prüfen Sie die Ergebnisse, wenn der Job beendet ist.
- Zwischendatensätze müssen bei der ersten Nutzung über JoSuA erneut erstellt werden, da JoSuA nicht auf **data_remote** zugreifen kann. Planen Sie dies bitte rechtzeitig mit ein, falls die Erstellung dieser Zwischendaten sehr zeitaufwändig ist.
- Ado-files, die Sie im Remote-Desktop über „fdzinstall" installiert haben, müssen in JoSuA erneut einmalig installiert werden.
- Falls Sie Skripte zwischen der Gastaufenthalts- und Remote-Desktop-Oberfläche transferieren wollen, laden Sie diese in JoSuA über die Funktion „Transfers" hoch. Sie stehen Ihnen nach erfolgter Prüfung zum Download bereit.
- Der Export von Ergebnissen erfolgt wie gewohnt über den Presentation / Publication Modus von JoSuA (siehe Kapitel 4). Bitte prüfen Sie auch beim Arbeiten im Remote Desktop zunächst immer, ob Ihre Programme auch vollständig lauffähig sind, indem Sie einen Lauf im Internal Use Modus zwischenschalten.

7 Ergebnisübermittlung

Ergebnisse werden fast ausschließlich im Rahmen von Jobs im Presentation / Publication Modus in JoSuA überprüft und freigegeben. Ausnahmen stellen lediglich generierte aggregierte Datensätze dar. Die Datenschutzprüfung für Auswertungen aus dem Presentation / Publication Modus kann bis zu **5 Arbeitstage** in Anspruch nehmen.

Im Folgenden werden einige technische und organisatorische Hinweise zur Ergebnisübermittlung gegeben. Weitere Informationen zu Begrenzungen hinsichtlich der Zahl und Größe von Ergebnisdateien pro Job in JoSuA finden Sie in Abschnitt 10.8. Eine ausführliche Anleitung finden Sie in den Richtlinien.

7.1 Datenschutzrichtlinien

Ergebnisse, die vom FDZ geprüft werden sollen, dürfen keine Angaben enthalten, die auf weniger als **20 Beobachtungen** beruhen. Ausnahmen, für die das FDZ eine manuelle Löschung einzelner Werte vornimmt, sind in den Skripten zu begründen. Die Nutzenden des FDZ sind daher angehalten, Ihre Ergebnisse, die vom FDZ geprüft und freigegeben werden sollen, so zu gestalten, dass diese schnell nachvollziehbar und keine manuellen Löschungen erforderlich sind. Hierfür müssen sich die Nutzenden an die Richtlinien halten. Das FDZ unterstützt Sie hierbei mit Stata-Templates, die Beispiele zur Dokumentation und datenschutzkonformen Programmierung enthalten.

Bitte lesen Sie die Richtlinien, bevor Sie mit den Analysen in der FDZ-Infrastruktur beginnen. Nutzen Sie die FDZ-Templates zur Gestaltung Ihrer Programmdateien.

7.2 Zulässige Ausgabeformate

- Log-Files:
  - Stata-Ergebnisdateien müssen entweder die Endung „.log" oder „.txt" tragen.
  - Ergebnisse, die in separate Dateien außerhalb der von Stata angelegten Log-Files geschrieben wurden, können nicht bereitgestellt werden. Sie müssen wieder in das Log-File eingebunden werden (siehe Beispiel in den FDZ-Templates). Zur

Datenschutzprüfung müssen die Ergebnisse unmittelbar vor dem eingebundenen Text nochmals in lesbarer Form mit Hilfe eines Stata-Befehls dargestellt werden.

- Grafiken:
  - Stata-Grafiken (.gph) können nur bereitgestellt werden, wenn sie mit der Option **asis** erzeugt wurden. Die Formate eps, svg und png sind ebenfalls zulässig. Hierbei ist zu beachten, dass die Formate eps und svg Vektorgrafiken erzeugen, die beliebig groß skalierbar sind und somit für besonders detaillierte Grafiken zu favorisieren sind.
  - Im Internal Use Modus kann nur das Grafikformat **png** dargestellt werden.
  - Speichern Sie die Grafiken im Verzeichnis **log** ab.
  - Setzen Sie die Richtlinien bei der Erstellung von Grafiken um.
  - Beispiele zur Ausgabe von Grafiken, die mit den FDZ-Richtlinien konform sind, finden Sie im FDZ-Template **jd05_graphs.do**.
- Generierte aggregierte Datensätze:
  - Falls Sie aggregierte Datensätze als Ergebnisdatei erhalten möchten, sprechen Sie die Vorgehensweise sowie Inhalt und Aufbau der Datensätze bitte vorher mit uns ab.
  - Die Generierung der aggregierten Daten sollte möglichst während eines Gastaufenthalts erfolgen, damit Sie das Ergebnis vorab im Detail prüfen können. Die Programme müssen anschließend nochmals mit JoSuA gestartet werden. Bitte verwenden Sie dafür den Presentation / Publication Modus und vermerken Sie im Kommentarfeld, dass Sie einen generierten Datensatz erhalten möchten und den entsprechenden Dateinamen. Speichern Sie diesen Datensatz immer im Unterverzeichnis **data**.
  - Für die Dokumentation der aggregierten Datensätze gelten die Richtlinien. Weitere Hinweise und Beispiele zur Ausgabe von Aggregaten, die mit den FDZ-Richtlinien konform sind, finden Sie im FDZ-Template **jd06_export_aggregate.do**.
  - Für jedes Projekt kann **nur einmal** ein solcher Datensatz weitergegeben werden. Dies gilt auch, wenn der aggregierte Datensatz in einem anderen FDZ-Projektverzeichnis zur Zuspielung bereitgestellt werden soll. Die Übermittlung des Datensatzes erfolgt per E-Mail. Transfers in andere Projektverzeichnisse übernimmt das FDZ.

8 Publikation der Ergebnisse

In Publikationen sind Daten und Dokumentationen entsprechend den Vorgaben auf der Website des FDZ zu zitieren:

- Seit 2018 erhalten alle FDZ-Datenprodukte eine DOI (Digital Object Identifier). Die DOI sowie genaue Vorgaben zur Zitation finden Sie auf der jeweiligen Seite des Datenprodukts auf unserer Website.
- Für ältere Datenprodukte ohne DOI finden Sie auf der FDZ-Website die Zitiervorgaben.

Senden Sie bitte ein Belegexemplar an das FDZ, idealerweise in digitaler Form.

9 Sonderregelungen der Softwarenutzung

- Standardmäßig bietet das FDZ für seine Nutzenden im Gastaufenthalt und in der Datenfernverarbeitung nur Stata als Analysesoftware an. In Ausnahmefällen kann andere Software (z. B. auf GNU-Octave oder R) vom FDZ zur Verfügung gestellt werden. Setzen Sie sich dazu bitte frühzeitig mit dem FDZ in Verbindung.
- Bei Verwendung anderer Software als Stata gilt grundsätzlich, dass alle Datenaufbereitungs- und Analyseschritte, die mit Stata machbar sind, auch mit Stata durchgeführt werden müssen.
- Verwenden Sie auch für die Sondersoftware-Auswertungen Master- und Log-Dateien. Die entsprechenden Programmiervorlagen erhalten Sie auf Anfrage vom FDZ.
- Bei Verwendung anderer Software als Stata gelten ebenso die „Richtlinien zur Gestaltung der Programmdateien" (siehe Kapitel 3). Ausgenommen davon ist lediglich die Regel, dass keine manuellen Löschungen durch das FDZ-Team vorgenommen werden. Da hierbei eine Vorab-Prüfung der Ergebnisse durch die Nutzenden nicht möglich ist, nimmt das FDZ-Team ausnahmsweise manuelle Löschungen vor.
- Da die Nutzung von Sondersoftware über JoSuA nicht möglich ist, müssen Skripte für die **kontrollierte Datenfernverarbeitung per E-Mail** an iab.fdz@iab.de verschickt werden. Hierbei müssen folgende Regelungen beachtet werden:
  - Zur Bearbeitung Ihrer Aufträge benötigen wir in jeder E-Mail Angaben zum Projekt in der Betreffzeile (siehe Kasten 1).
  - Neu eingeschickte Dateien überschreiben bereits vorhandene, wenn sie den gleichen Dateinamen haben.
  - Unter Sonstiges bitte schon in der Betreffzeile auf Dinge hinweisen, die bei der Bearbeitung der Programme zu beachten sind (z. B. „GNU-Octave").
- Wir bitten um Verständnis dafür, dass wir Aufträge unbearbeitet zurücksenden, wenn relevante Informationen in der Betreffzeile fehlen oder die Formvorgaben nicht erfüllt sind.
  - Senden Sie für ein Projekt keine neuen Programme ein, solange Sie die Ergebnisse der letzten kontrollierten Datenfernverarbeitung noch nicht erhalten haben. Nutzen Sie in dieser Zeit auch nicht JoSuA, um weitere Ergebnisse im Presentation / Publication Modus zu erzeugen.

- Vor einem Gastaufenthalt:
  - Programmdateien, die Sie im Gastaufenthalt verwenden wollen, können über „Transfers" in JoSuA hochgeladen werden.
- Nach einem Gastaufenthalt:
  - Nach einem Gastaufenthalt werden nicht die während des Aufenthalts erzeugten Ergebnisse herausgegeben, sondern die entsprechenden Skripte werden vom FDZ-Team nochmals gestartet. Sorgen Sie daher dafür, dass die entsprechende Master-Datei und die anderen Programmdateien vollständig lauffähig sind.
  - Wenn Sie nach Abschluss Ihres Gastaufenthalts Ergebnisse erhalten wollen, schicken Sie eine E-Mail an das FDZ-Postfach (siehe Kasten 1). Schreiben Sie zusätzlich „*nach GA*" in die Betreffzeile und geben Sie in der E-Mail das Datum des Gastaufenthaltes sowie die Ergebnis-Dateien an, die wir Ihnen schicken sollen.
  - Wir starten die Master-Datei und Sie erhalten die geprüften Ergebnisse, die in diesem Programmdurchlauf erzeugt wurden.
  - Skripte, die Sie im Gastaufenthalt erstellt oder bearbeitet haben, können über „Transfers" in JoSuA hochgeladen werden.
- Im Remote-Desktop kann das Arbeiten mit Sondersoftware **nicht** angeboten werden. Im Remote-Desktop erzeugte Ergebnisse müssen regelmäßig in JoSuA validiert werden (siehe Kapitel 6.3), JoSuA unterstützt jedoch keine Sondersoftware.

Kasten 1: Beispiele für Betreffzeile der E-Mail bei Datenfernverarbeitung mit Sondersoftware

Projektnummer – max. Speicherbedarf – erwartete Laufzeit – Datum Vertragsende – Sonstiges
(yyyy-mm-dd = Datum im Format Jahr-Monat-Tag)

fdz###, # GB, # Stunden/Tage, yyyy-mm-dd, [Sonstiges]

Beispiele: > > fdz0041, 3 GB, 14 h, 2025-12-31, R > > fdz0041, 3 GB, 14 h, 2025-12-31, nach GA am 2025-01-30, GNU-Octave

10 Bedienungsanleitung JoSuA

In den folgenden Kapiteln erfahren Sie, wie Sie sich in JoSuA einloggen, wie man durch die JoSuA-Webschnittstelle navigiert, Aufträge einsendet und anschließend auf die Ergebnisse zugreift. Darüber hinaus finden Sie Hinweise, was Sie bei möglichen Problemen mit JoSuA tun können.

Aktuelle Antworten auf häufige Fragen zu JoSuA finden Sie auch auf der JoSuA-Seite der FDZ-Website. Dort finden Sie auch ein Video-Tutorial, dass Ihnen den Einstieg zur Nutzung von JoSuA erleichtern soll.

10.1 Login

- Rufen Sie die Internetseite https://josua.iab.de/ auf. JoSuA kann nicht mit dem Internet Explorer verwendet werden.
- Geben Sie im Login-Fenster die Zugangsinformationen ein, die Sie vom FDZ erhalten haben (siehe Kapitel 4.1). Jeder Nutzer bzw. jede Nutzerin erhält nach Abschluss eines Nutzungsvertrags eine Kennung für JoSuA. Sollten mehrere Projektverzeichnisse einer Person zugeschrieben sein, gilt diese Kennung für alle Projektzugänge.[1]
- Falls Sie noch keine Zugangsdaten erhalten haben oder diese vergessen haben, setzen Sie sich mit dem FDZ in Verbindung (iab.fdz@iab.de).
- Die Weitergabe des Passworts an Dritte ist nicht erlaubt, auch nicht innerhalb eines Projekts.

10.2 Start

Nach der erfolgreichen Anmeldung erscheint die Nutzeroberfläche. Das Hauptmenü besteht aus mehreren Reitern: *Home*, *Projects*, *Submission* und *Transfers*. Diese werden in den folgenden Kapiteln näher beschrieben. Zu Beginn ist der Reiter *Home* ausgewählt. In der oberen rechten Ecke können Sie Ihr Passwort ändern oder sich ausloggen.

[1] Bei mehreren Nutzenden pro Projekt erhält jede/r Nutzende eigene Zugangsdaten.

10.3 Home

Der Reiter Home, der in Abbildung 2 dargestellt ist, zeigt eine Liste mit allen aktuellen Aufträgen, deren Status und, ob der Job erfolgreich durchgelaufen ist.

Folgende Statusinformationen, die den Stand des Jobs im Bearbeitungsprozess darstellen, sind möglich:

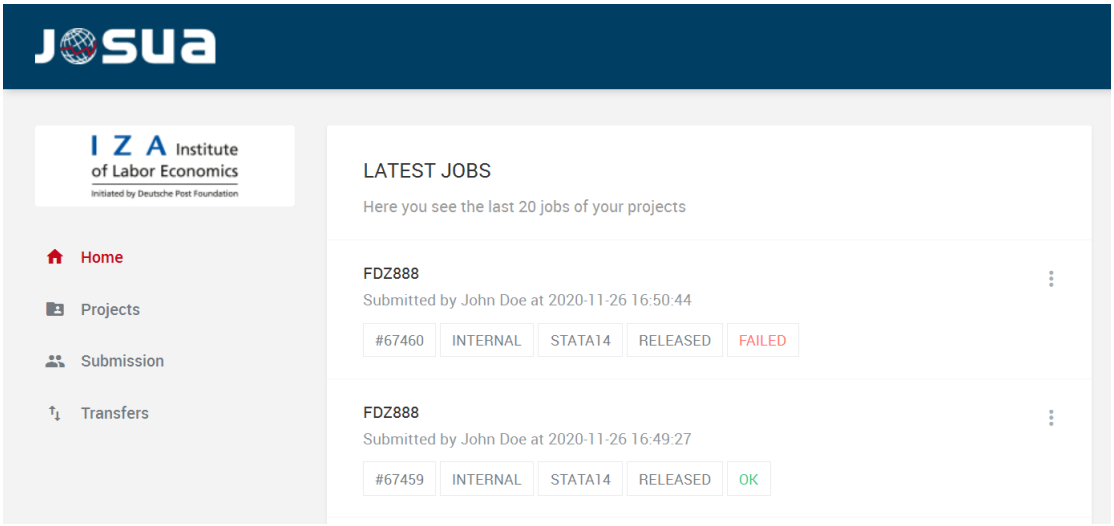
- „running": Der Job wird gerade ausgeführt.
- „censoring": Der Job befindet sich in der Zensierung (Presentation / Publication Modus) bzw. es werden die png-Dateien erzeugt (Internal Use Modus – siehe auch 10.10.2).
- „released": Der Job wurde freigegeben.

Folgende weitere Statusangaben können im letzten Feld der Übersicht erscheinen:

- „success": Der Job ist erfolgreich durchgelaufen.
- „failed": Der Job ist abgebrochen. In Ausnahmefällen erscheint der Status „failed", obwohl der Job erfolgreich durchgelaufen ist.
- „aborted": Der Job wurde von den Nutzenden selbst vorzeitig gestoppt.

Klicken Sie auf die drei Punkte auf der rechten Seite, um Details des ausgewählten Auftrags anzusehen.

Abbildung 2: Reiter *Home*

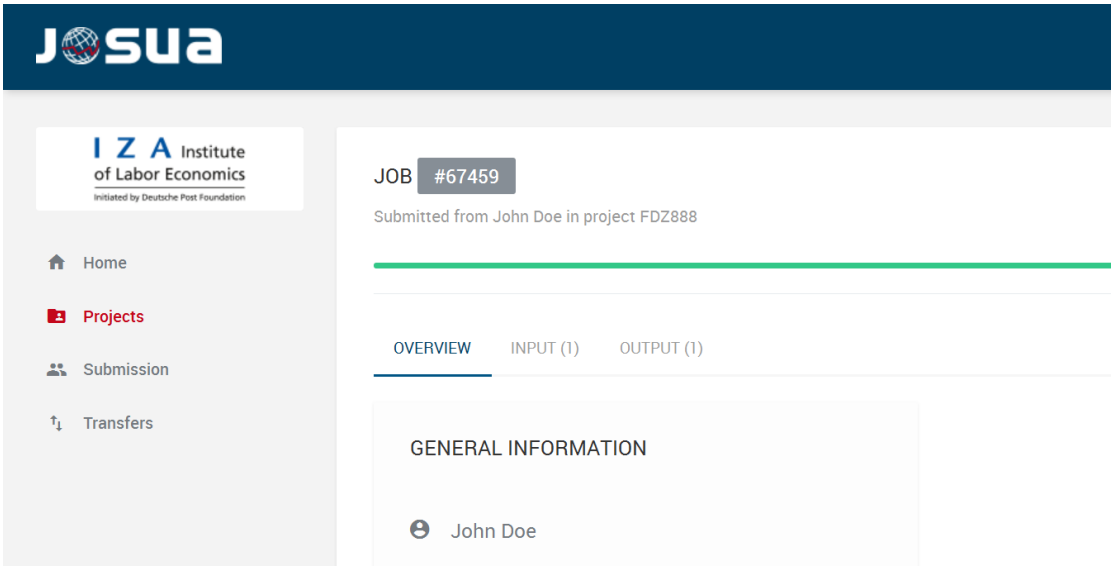


Neben einigen allgemeinen Informationen zu Ihrem Auftrag finden Sie in der Detailansicht (siehe Abbildung 3) unter *Input* alle hochgeladenen Dateien und unter *Output* alle verfügbaren

Ergebnisdateien. Derzeit können pro Auftrag maximal **20 Ergebnisdateien** generiert werden. Dabei werden alle Log-Files (inkl. einer etwaigen Dateiliste) und Grafiken mitgezählt. Log-Files werden automatisch abgeschnitten, wenn eine Dateigröße von **2 MB pro Datei oder alle Ergebnisdateien des Jobs zusammen 8 MB** überschreiten. Die Darstellung der Ergebnisse ist abhängig vom Einsendemodus.

- Input (Hochgeladene Dateien):
  - Alle hochgeladenen Skripte für diesen Auftrag werden hier aufgelistet.
- Output (Ergebnisdateien):
  - Internal Use: Sie können die Ergebnisse Ihres Auftrags per Klick auf das entsprechende Output-File (Log-File oder Grafik) oder das Icon auf der rechten Seite einsehen. Navigieren Sie mittels Klicken auf die Buttons rechts unten durch die Vorschaufenster. Bitte beachten Sie, dass die Ergebnisse nur solange zugänglich sind bis im gleichen Projekt ein neuer Auftrag im Presentation / Publication Modus hochgeladen wird.
  - Presentation / Publication: Nach der manuellen Datenschutzprüfung können alle geprüften Ergebnisse heruntergeladen werden. Der Link erscheint erst nach erfolgter Datenschutzprüfung. Um Dateien herunterzuladen, klicken Sie auf den Dateinamen oder auf den Pfeil auf der rechten Seite und speichern die Datei auf Ihrem Computer.

Abbildung 3: Detailansicht zum ausgewählten Auftrag



10.4 Projects: Projektübersicht

Im Reiter *Projects* (Projekte) sind alle Ihre laufenden Projekte aufgelistet (siehe Abbildung 4).

Durch einen Klick auf "*View*" gelangen Sie zu den Projektdetails des entsprechenden Projekts. In diesem Menü finden Sie drei Reiter:

- **Jobs**: Hier finden Sie eine Liste der Aufträge aller Nutzenden des gewählten Projektes (siehe Abbildung 5). Zusätzlich zu generellen Informationen können Sie durch Anklicken der drei Punkte auf der rechten Seite dieselben Optionen aufrufen, die Sie auch unter *Home* finden (siehe Kapitel 10.3).
- **Resources**: Diese Reiter dient der Übermittlung und Speicherung von zusätzlichen Befehlsbibliotheken, z. B. Stata-Ado-Paketen. Laden Sie hier nur Ado-Files für Ihr Projekt hoch (siehe Abbildung 6), wenn diese nicht über den Befehl `fdzinstall` aus den am FDZ verfügbaren Paketen aus dem SSC-Archive installiert werden können (siehe Kapitel 2.2.1)! Alle Nutzenden eines Projekts können Dateien hinzufügen, die alle autorisierten Mitglieder für ihre Aufträge nutzen können. Falls dort schon von anderen Mitgliedern hochgeladene Dateien sind, können Sie diese hier sehen. Um neue Dateien hinzuzufügen klicken Sie auf "+" und ziehen Sie die gewünschte Datei in das Fenster oder klicken Sie in das Fenster, um die Datei von Ihrer Festplatte hochzuladen. Es ist nur möglich, einzelne Dateien hochzuladen. Falls bereits eine Datei mit gleichem Namen existiert, wird diese überschrieben (v.a. zu beachten bei sich ändernden Versionen von Ado-Files). Bitte beachten Sie, dass über *Resources* ausschließlich reine Textdateien hochgeladen werden können. Das Hochladen von anderen Formaten (wie beispielsweise .mlib oder .plugin) ist nicht möglich (siehe Kapitel 2.2.1). Nach dem Hochladen erscheint das Ado-File in der Liste mit der Statusangabe „*new*". Sobald das Ado-File nach einer Datenschutzprüfung freigegeben wurde, wechselt der Status auf „*accepted*" und es steht dauerhaft im Projekt zur Nutzung innerhalb von JoSuA zur Verfügung. Wurde das Ado-File nicht freigegeben, erhält es den Status „*dropped*". Um über RESOURCES hochgeladene Ado-Pakete auch für Gastaufenthalte verfügbar zu machen, müssen sie während des Gastaufenthalts einmalig heruntergeladen und im **prog**-Verzeichnis gespeichert werden.
- **Transfers**: Hier sind alle Skripte aufgelistet, die Sie vor oder nach einem Gastaufenthalt zum Transfer für dieses Projekt hochgeladen haben (siehe Abbildung 7). Nach Prüfung der Skripte durch das FDZ (*censored*) können Sie die Dateien für die weitere Bearbeitung zu Hause oder über Datenfernverarbeitung durch einen Klick auf den Dateinamen oder den Pfeil auf der rechten Seite herunterladen. Sie können durch einen Klick auf „+" auch neue Skripte hochladen. Ausführliche Informationen zu Transfers finden Sie in Kapitel 10.7.

Abbildung 4: Projektansicht

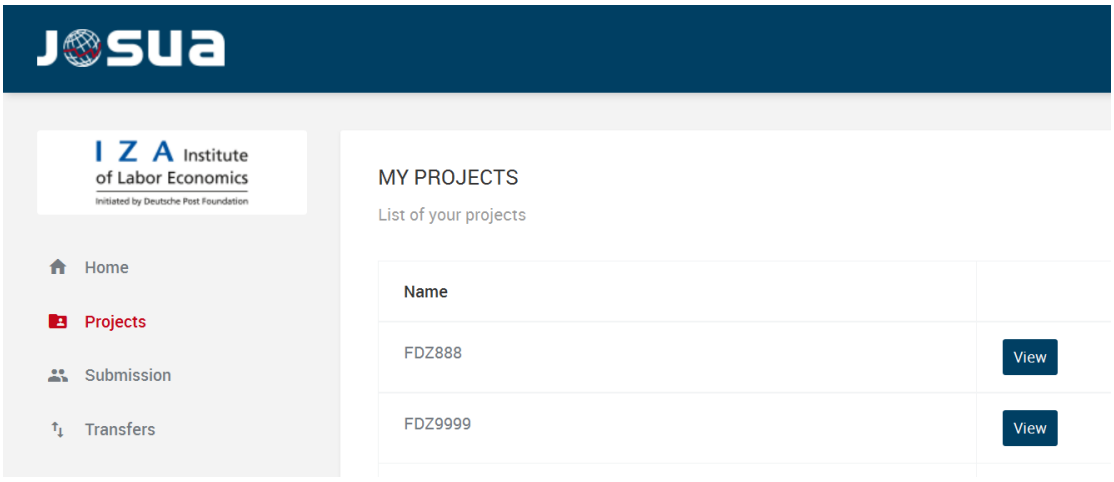


Abbildung 5: Projektinformationen

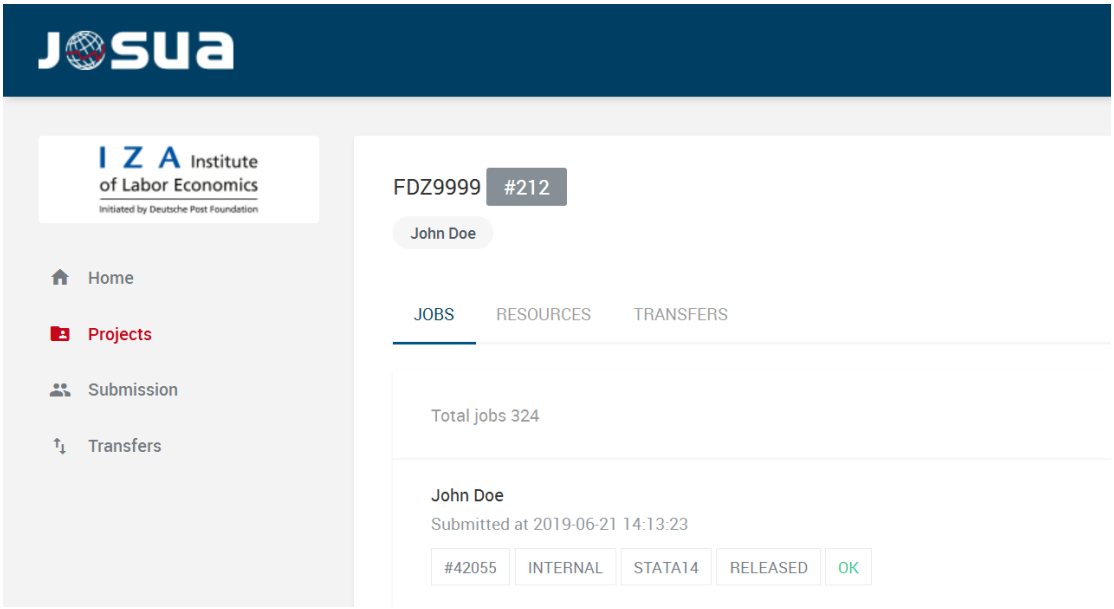


Abbildung 6: Resources (Ado-Dateien)

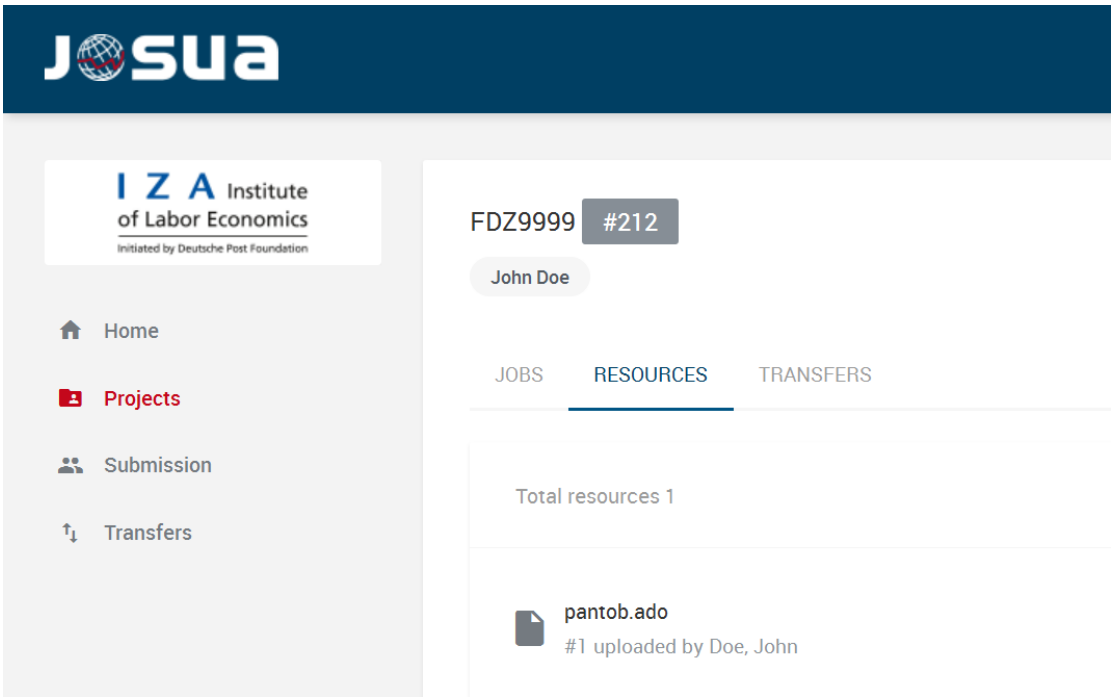


Abbildung 7: Transfers

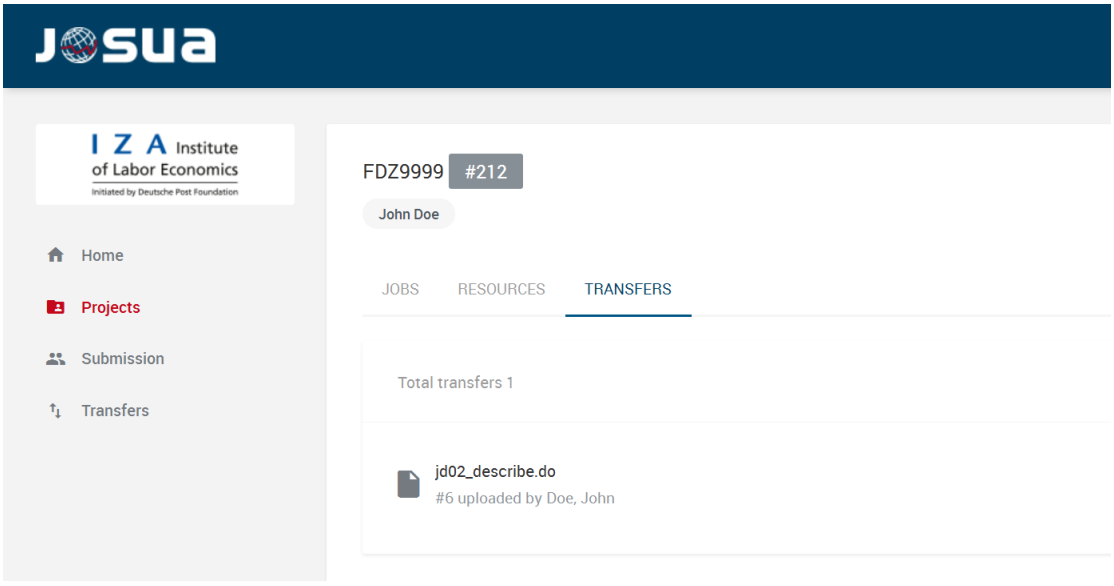
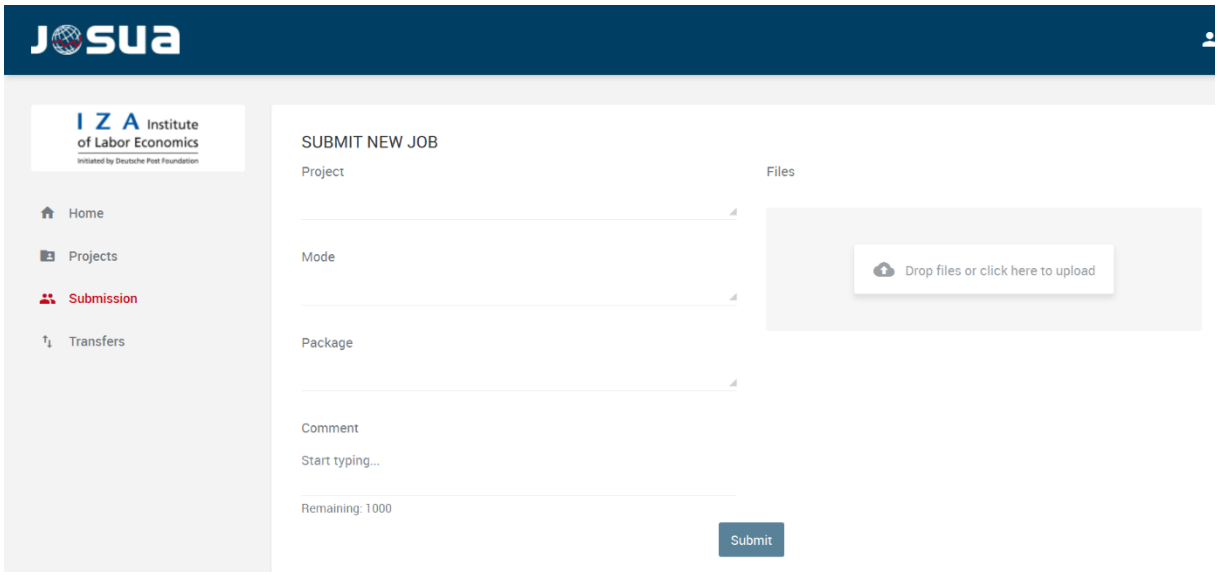


Abbildung 8: Hochladen von Aufträgen



10.5 Submission: Hochladen von Aufträgen

Ein Auftrag zur Datenfernverarbeitung wird übermittelt, indem die Skripte im Reiter *Submission* (Hochladen von Aufträgen) hochgeladen werden (siehe Abbildung 8). Es muss eine Master-Datei namens `master.do` enthalten sein. Bei jedem Auftrag müssen alle auszuführenden Skripte hochgeladen werden. Bitte verwenden Sie bei der Benennung der Skript-Dateien nur folgende Zeichen: 0-9, a-z, A-Z, -,_.

Auf der linken Seite wählen Sie im ersten Drop-Down-Menü bitte das Projekt aus, in welchem Sie den Job hochladen wollen (z. B. fdz2020). Im zweiten Drop-Down-Menü können Sie auswählen, ob der Auftrag im Internal Use Modus oder Presentation / Publication Modus eingeschickt werden soll (siehe Kapitel 4). Im dritten Drop-Down-Menü wählen Sie bitte das Software-Paket, das genutzt werden soll.

Im darunterliegenden Feld können Sie einen Kommentar eingeben, der relevante Details oder Erklärungen enthält. Verzichten Sie auf Umlaute und Sonderzeichen und verwenden Sie maximal 1.000 Zeichen.

Ziehen Sie anschließend Ihre Skripte in das Fenster auf der rechten Seite oder klicken Sie in das Fenster, um Skripte von Ihrer Festplatte in JoSuA hochzuladen. Bitte laden Sie pro Auftrag einschließlich der Master-Datei maximal 20 Skripte hoch. Die hochgeladenen Dateien werden aufgelistet. Sie können die Liste prüfen und einzelne Dateien entfernen. Wenn Sie mit dem Hochladen Ihrer Skripte fertig sind, klicken Sie auf *Submit* (Auftrag übermitteln), um Ihren Auftrag zu übersenden.

Sobald Sie einen Job im Presentation / Publication Modus hochgeladen haben, können Sie bis zu dessen Freigabe keine weiteren Jobs im Presentation / Publication Modus hochladen. Im Internal Use Modus ist es theoretisch möglich, mehrere Jobs parallel laufen zu lassen. Ebenso können, während ein Job im Presentation / Publication Modus läuft, weitere Jobs im Internal Use Modus gestartet werden. Das FDZ empfiehlt, diese Möglichkeiten nur in Ausnahmefällen zu nutzen, wenn z. B. ein Job im Presentation / Publication Modus mehrere Tage läuft. Bitte laden Sie im Normalfall erst dann einen neuen Job hoch, wenn der vorherige Job durchgelaufen ist.

10.6 Stoppen von Jobs

Falls ein laufender Job gestoppt werden soll, können Sie dies selbst vornehmen. Im Reiter „*Home*" sind Ihre Jobs aufgelistet. Beim entsprechenden Job befindet sich rechts eine Multifunktionsschaltfläche mit drei übereinander angeordneten Punkten. Damit öffnet sich das Kontextmenü des Jobs. Dort befindet sich wiederum rechts oben eine Multifunktionsschaltfläche mit drei übereinander angeordneten Punkten. Hinter dieser verbirgt sich die Funktion „*Stop Job*". Wenn Sie diese Schaltfläche anklicken und den Sicherheitshinweis bestätigen, wird der Job gestoppt. Der Job hat nun den Status „*aborted*". Jobs, die im Presentation / Publication Modus hochladen wurden, werden damit jedoch nicht automatisch gelöscht, sondern liegen zur Prüfung und Freigabe durch das FDZ bereit. Das FDZ löscht diese Jobs anschließend manuell. Dies bedeutet, dass bis zur endgültigen Löschung keine weiteren Jobs im Presentation / Publication Modus eingereicht werden können.

10.7 Transfers: Übermittlung von Skripten vor oder nach dem Gastaufenthalt

Über den Reiter *Transfers* können Sie zu Hause vorbereitete Skripte für Ihren Gastaufenthalt hochladen (siehe Abbildung 9). Dies gilt neben Stata-Dofiles auch für Skripte anderer Programme wie z.B. R oder Octave. Sollten Sie nach Ihrem Gastaufenthalt die während des Gastaufenthalts erstellten Skripte zur weiteren Bearbeitung zu Hause oder über Datenfernverarbeitung benötigen, laden Sie diese bitte auch über die Funktion „*Transfers*" hoch. Laden Sie über „*Transfers*" keine Ado-Files oder andere Programm-Pakete hoch (siehe 2.2.1)

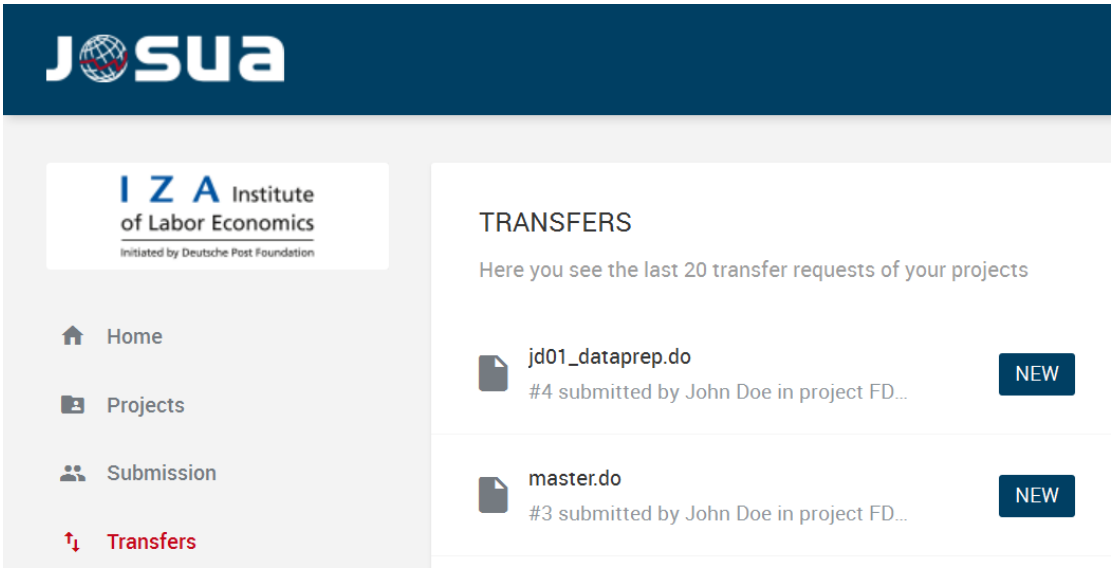
Ziehen Sie das gewünschte Skript hierzu in das Fenster auf der rechten Seite oder klicken Sie in dieses Fenster, um das Skript von Ihrer Festplatte hochzuladen. Alternativ können Sie auch über den Reiter *Projects* -> View -> TRANSFERS -> + gehen. JoSuA akzeptiert hier – im Gegensatz zum Hochladen von Dateien bei *„Submission"* - nur das Hochladen / Drag & Drop **einzelner Dateien.** Es ist aber möglich, mehrere Dateien in einer Archiv-Datei (.zip, .7z) hochzuladen.

Nach dem Hochladen erscheint das Skript in der Liste mit der Statusangabe „new". Sobald das Skript nach einer Datenschutzprüfung freigegeben wurde, wechselt der Status auf „accepted". Die Skripte stehen nun zum Herunterladen zur Verfügung. Sollten durch das FDZ Löschungen in den Skripten vorgenommen werden, erhalten die Dateien den Status „censored". In der Regel finden Sie einen Hinweis zu den Löschungen am Anfang des Skripts. Vereinzelt erhalten Dateien jedoch auch den Status „censored", obwohl keine Löschungen vorgenommen wurden. Dies ist technisch bedingt.

Unter dem Reiter Transfers sehen Sie alle bereits transferierten Skripte. Zudem haben Sie hier die Möglichkeit, die durch das FDZ-Team freigegebenen Skripte herunterzuladen. Klicken Sie hierzu auf den Skript-Namen oder den Pfeil auf der rechten Seite.

Damit eine Archiv-Datei freigegeben werden kann, müssen alle darin enthaltenen Dateien datenschutzrechtlich unbedenklich sein, ansonsten wird sie komplett gelöscht. Bitte beachten Sie, dass beim Erstellen einer Archiv-Datei an einem Rechner mit macOS automatisch Systemfiles/Metadaten erzeugt werden. Diese müssen von Ihnen gelöscht werden, bevor Sie die Archiv-Datei in JoSuA hochladen. Eine Löschung der Systemfiles, die für Sie selbst nicht sichtbar sind, kann mit folgendem Befehl über die Kommandozeile gelöscht werden: `zip -r data.zip . -x ".*" -x "__MACOSX"`.

Abbildung 9: Übermittlung von Do-Files



10.8 Technische Restriktionen in JoSuA

Bei der Nutzung von JoSuA gibt es einige Beschränkungen, die im Folgenden nochmals zusammengefasst dargestellt werden:

- Unter „*Submission*" können einschließlich der Master-Datei **maximal 20 Skripte** pro Job hochgeladen werden.
- Es können **maximal 20 Ergebnisdateien** pro Job ausgegeben werden. Hierzu zählt auch jede einzelne Grafik oder eine etwaige Dateiliste.[2]
- Maximale Größe einer Ergebnisdatei: **2 MB.** Sollte ein Log-File zu groß sein, wird die Datei abgeschnitten und ein entsprechender Hinweis in das Log-File geschrieben. Sollte eine Grafik-Datei zu groß sein, wird diese gelöscht und die Löschung in einem Log-File „lost-files.log" dokumentiert.
- Maximale Größe aller Ergebnisdateien pro Job zusammen: **8 MB**
- Pro Projekt ist das Unterverzeichnis **data** auf die Größe von **30 GB** beschränkt.
- Es stehen maximal **30 GB Arbeitsspeicher** zur Verfügung. Jobs, die zur Laufzeit mehr Arbeitsspeicher erfordern, brechen automatisch ab.
- Pro Tag kann nur eine begrenze Anzahl an Jobs pro Projekt im Internal Use Modus bzw. im Presentation / Publication Modus hochgeladen werden. Die aktuell gültigen Limits finden Sie auf der FDZ-Website.

10.9 Hilfe bei Job-Abbrüchen in JoSuA

Sollte Ihr Job in JoSuA mit dem Status „*failed*" abbrechen, prüfen Sie bitte zunächst eigenständig folgende Aspekte:

- Ist ein Job im Presentation / Publication Modus abgebrochen, beachten Sie bitte die Fehlermeldung im letzten Log-File bzw. einem etwaigen master.log sowie einen etwaigen User Comment des FDZ, um dem Problem auf den Grund zu gehen.
- Überprüfen Sie Ihre Programme auf Programmierfehler.
- Stellen Sie sicher, dass Sie alle Vorgaben zur Datennutzung – insbesondere die Pfadangaben betreffend – eingehalten haben.
- Als Vorschlag, hier wird empfohlen in den FAQ nachzusehen, es könnte hier zusätzlich auf die FAQ verlinkt werden.

Erst wenn diese Maßnahmen zu keiner Lösung führen, wenden Sie sich bitte an das FDZ. Übermitteln Sie dazu bitte folgende Informationen:

- Projektnummer, Jobnummer in JoSuA
- Möglichst detaillierte Darstellung des Problems (z. B. Wortlaut der Fehlermeldung)

[2] Im Internal Use Modus bezieht sich die Begrenzung auf die Anzahl der Dateien vor der Umwandlung in die Vorschaubilder.

Übermitteln Sie bitte **NIEMALS** Ergebnisse, die im Internal Use Modus erstellt wurden oder sonstigen nicht vom FDZ-Team freigegebenen Output! Die Übermittlung von nicht freigegebenen Ergebnissen – auch an das FDZ – stellt eine Vertragsverletzung dar.

In Ausnahmefällen erscheint der Status „*failed*", obwohl der Job erfolgreich durchgelaufen ist.

10.10 Fehlerbehebung in JoSuA

Im Folgenden werden häufige Fehler und deren Behebung beschrieben.

Bei Fragen kontaktieren Sie bitte das Forschungsdatenzentrum (iab.fdz@iab.de). Bitte geben Sie immer Ihre FDZ-Projektnummer (**fdzXXXX**) und ggf. die **Job-ID** aus JoSuA an.

10.10.1 Dateien können nicht hochgeladen werden

Unter "Transfers" und "RESOURCES" können Dateien nur einzeln hochgeladen werden (im Unterschied zu "Submission"). Beim Transferieren von Skripten vor / nach Gastaufenthalt ("Transfers") können Sie die einzelnen Dateien aber vorher packen (.zip, .7z).

10.10.2 Freigabe der Ergebnisse im Internal Use Modus dauert ungewöhnlich lang (Status „censoring")

Üblicherweise erfolgt die Freigabe von Ergebnissen im Internal Use Modus innerhalb weniger Minuten. Bei hoher Beanspruchung von JoSuA kann es jedoch auch einige Stunden dauern. Bitte warten Sie mindestens 6 Stunden ab, bevor Sie das FDZ kontaktieren.

10.10.3 Ergebnisdateien fehlen

Nach Freigabe der Ergebnisse sind nicht alle im Job erzeugten Ergebnisdateien unter „*Output*" verfügbar. Hierfür kann es unterschiedliche Gründe geben:

- Es wurden ganze Dateien im Rahmen der Datenschutzprüfung nicht freigegeben. In diesem Fall finden sich in der Regel Hinweise zu den Löschungen in den Kommentaren.
- Pro Job wurden mehr als 20 Ergebnisdateien erzeugt. Hierzu zählen auch jede einzelne Grafik und eine etwaige Dateiliste. JoSuA gibt dann nur maximal 20 (zufällig ausgewählte) Ergebnisdateien aus.

- Die maximale Größe von 2 MB pro Ergebnisdatei wurde überschritten. Sollte ein Log-File zu groß sein, wird die Datei abgeschnitten und ein entsprechender Hinweis in das Log-File geschrieben. Sollte eine Grafik-Datei zu groß sein, wird diese gelöscht und die Löschung in einem Log-File „lost-files.log" dokumentiert.
- Im Internal Use Modus: Es gab einen Fehler bei der Umwandlung in png-Dateien. Bitte kontaktieren Sie in diesem Fall das FDZ unter Angabe der Projektnummer und der Job-ID.

10.10.4 Fehlermeldung im Log-File "File not found"

Diese Fehlermeldung tritt auf, wenn der Dateipfad nicht mit den vorgesehenen Globals **$prog**, **$data**, **$orig** oder **$log** angegeben wurde. Verwenden Sie bei jeder Pfadangabe die genannten Globals, da JoSuA sonst die Dateien nicht finden kann. Bitte achten Sie außerdem darauf, keine Umlaute, Leerzeichen oder Sonderzeichen in Dateinamen zu benutzen.

10.10.5 Log-File master.log / Fehlermeldung "log-file already open"

JoSuA legt automatisch eine Log-Datei namens master.log an. Bitte legen Sie nicht selbst ein master.log an. Verwenden Sie `"cap log close"`, um das geöffnete Log-File zu schließen und legen Sie anschließend ein neues Log-File an. Hinweise zur Benennung von Skripten und Log-Files finden Sie in den Richtlinien.

10.10.6 Fehlermeldung "File not ok (suspicious mismatch between file suffix and actual file type)"

Diese Fehlermeldung erscheint beispielsweise, wenn das Log-File in Statas smcl-Format gespeichert wird. Bitte benutzen Sie nur Log-Files im reinen Text-Format. Dies erreichen Sie durch die Befehle `"$log\logfile.log", replace"` oder `"log using "$log\logfile"`, text replace".

10.10.7 Fehlermeldung "Error!"

Der Fehler tritt auf, wenn die Gesamtgröße aller generierten Output-Files eines Jobs größer als 8 MB ist. Bitte passen Sie Ihre Programme an, damit weniger Output generiert wird.

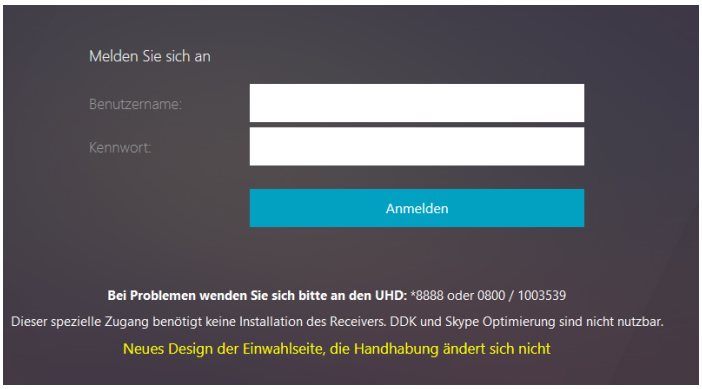
11 Bedienungsanleitung Remote-Desktop

In den folgenden Kapiteln erfahren Sie, wie Sie sich im Remote-Desktop-Zugang einloggen, wie Sie durch die Ansicht navigieren, sich abmelden bzw. den Zugang trennen und das Passwort ändern.

11.1 Login Remote-Desktop-Zugang

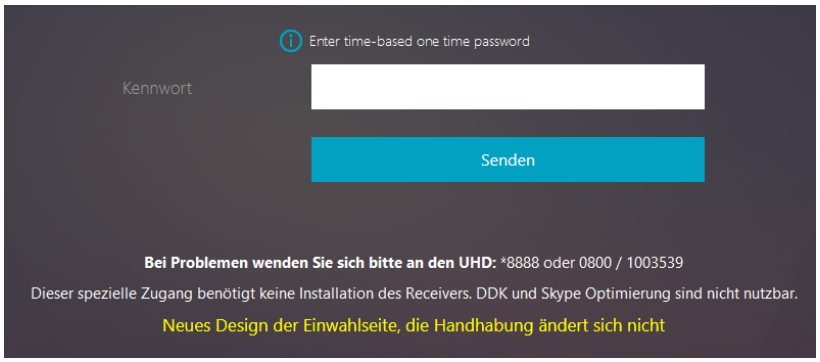
- Rufen Sie die Internetseite https://vdi.arbeitsagentur.de/ mit einem aktuellen Browser auf.
- Geben Sie im Login-Fenster (siehe Abbildung 10) in das Feld „Benutzername" bitte Ihren Remote-Desktop Benutzernamen mit dem Anhang ***@dst.baintern.de*** ein (z. B. für das Projekt fdz9999 und den Benutzer John Doe ist das voraussichtlich: fdz9999dj001@dst.baintern.de). In das Feld Kennwort tragen Sie bitte Ihr projektbezogenes Gastaufenthaltskennwort ein.
- Falls Sie noch keine Zugangsdaten erhalten haben oder diese vergessen haben, setzen Sie sich mit dem FDZ in Verbindung (iab.fdz@iab.de).
- Die Weitergabe des Passworts an Dritte ist nicht erlaubt, auch nicht innerhalb eines Projekts.

Abbildung 10: Login Remote Desktop Zugang



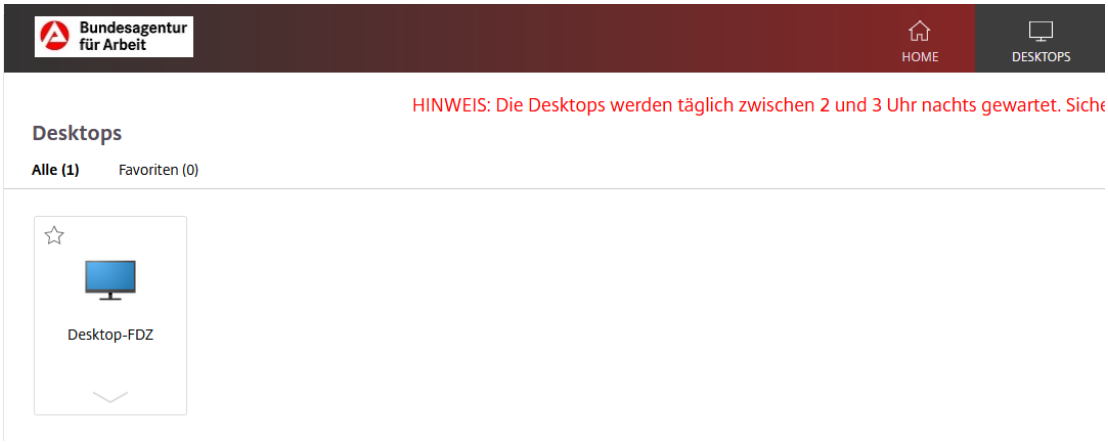
- Anschließend öffnet sich eine neue Seite (siehe Abbildung 11). In das Eingabefeld unter „Enter time-based one time password" tragen Sie bitte die mit Ihrem TOTP-Generator Ihres Smartphones aktuell generierte sechsstellige Sicherheitszahl ein.

Abbildung 11: Eingabe „Enter time-based one time password"



- Anschließend öffnet sich eine neue Seite (siehe Abbildung 12). Bitte klicken Sie das blaue Monitor Symbol am linken Rand an. Je nach Browser und dessen Einstellung müssen Sie danach eventuell noch eine automatisch heruntergeladene Datei per Doppelklick ausführen.

Abbildung 12: Verbindung zum Desktop-FDZ: Citrix Workspace



- Haben Sie bitte etwas Geduld, die Verbindung zum Server wird aufgebaut.

11.2 Zusätzliche Navigationselemente

Wenn die Verbindung aufgebaut ist, dann stehen Ihnen nun folgende Navigationselemente zur Verfügung:

- Hinter dem Halbkreis in der Mitte am oberen Rand verbirgt sich eine Multifunktionsschaltfläche (siehe Abbildung 13).

Abbildung 13: Multifunktionsschaltfläche: Zusätzliche Navigationselemente 1

- Durch Anklicken kann diese erweitert und wieder minimiert werden (siehe Abbildung 14).

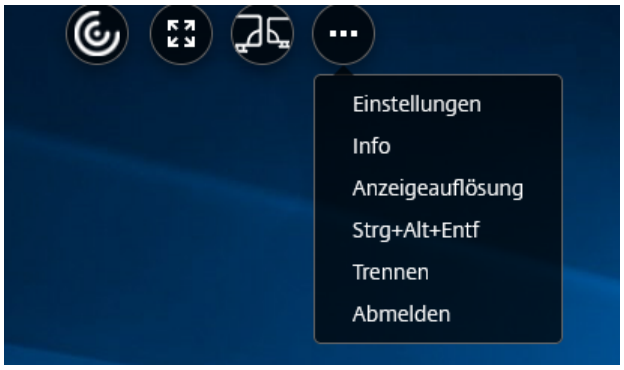
Abbildung 14: Erweiterung und Minimierung der Multifunktionsschaltfläche: Zusätzliche Navigationselemente 2

• Die Vollbildfunktion finden Sie hinter der zweiten Schaltfläche von links (siehe Abbildung 15).

Abbildung 15: Die Vollbildfunktion: Zusätzliche Navigationselemente 3

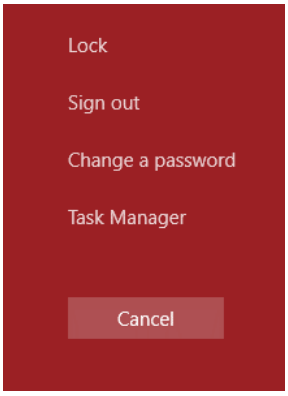
- Weitere Funktionen finden Sie hinter der Schaltfläche mit den drei Punkten (siehe Abbildung 16).

Abbildung 16: Weitere Funktionen: Zusätzliche Navigationselemente 4



- Wenn Sie Strg+Alt+Entf auf der Tastatur drücken, wirkt sich das nur auf Ihren Computer aus.
- Wenn Sie die Tastenkombination Strg+Alt+Entf hier auf dem Server ausführen möchten, verwenden Sie bitte die entsprechende Schaltfläche.
- Wenn Sie die Schaltfläche Strg+Alt+Entf betätigen, dann öffnet sich ein Menü (s. Abbildung 17), in welchem Sie:
  - Ihr Passwort ändern,
  - sich abmelden (Sign out) und
  - die Session sperren können.

Abbildung 17: Menü der Schaltfläche Strg+Alt+Entf: Zusätzliche Navigationselemente 5



11.3 Trennen oder Abmelden?

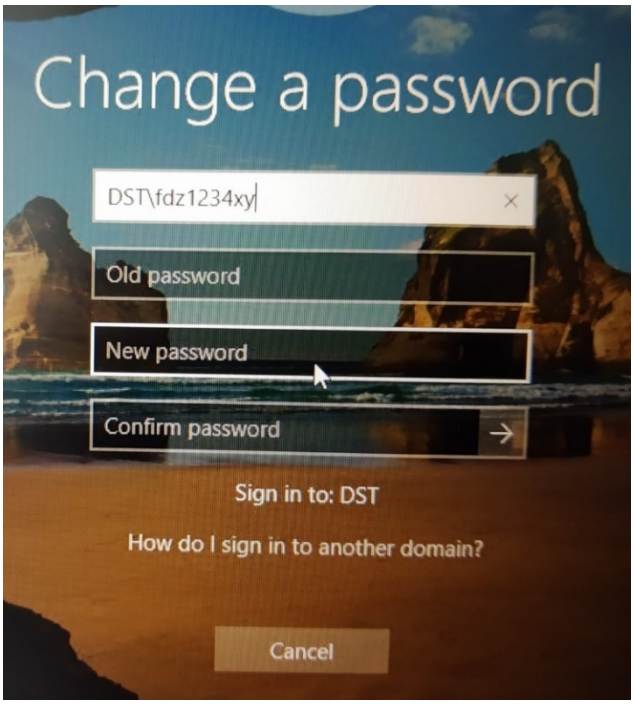
- Über die drei Punkte der Multifunktionsschaltfläche können Sie die Verbindung zum Server wieder trennen und sich abmelden (siehe Abbildung 16).
- Beim „Trennen" wird die Verbindung zum Server getrennt, aber alle Prozesse, die auf dem Server laufen (z.B. Stata-Auswertungen) werden fortgesetzt. Bitte nutzen Sie diese Variante, wenn Ihre Auswertungsprogramme noch laufen.
- Beim „Abmelden" wird Ihre Session heruntergefahren. Alle laufenden Prozesse werden gestoppt. Bitte melden Sie sich ab, wenn Ihre Stata-Auswertungsprogramme beendet sind, um nicht unnötig Ressourcen zu verbrauchen. Andere Forschende werden Ihnen den Performancegewinn danken!

11.4 Passwort ändern

Im Folgenden zeigen wir Ihnen, wie Sie Ihr Passwort ändern können. Hiermit ändern Sie aber nur das Passwort der Remote-Desktop-Kennung und nicht das Passwort Ihrer JoSuA-Kennung oder der Gastaufenthaltskennung!

- Melden Sie sich bitte zunächst in Ihrem Projekt an (siehe Abbildung 10 bis Abbildung 12).
- Öffnen Sie die Multifunktionsschaltfläche (oben in der Mitte), indem Sie das „Halbkreis-Symbol" anklicken (siehe Abbildung 13).
- Wichtig dabei ist, dass Sie die Vollbildansicht aktivieren! (siehe Abbildung 15).
- Klicken Sie, in der geöffneten Multifunktionsschaltfläche das „drei Punkte Symbol" an (siehe Abbildung 16) und klicken Sie die „Strg+Alt+Entf-Schaltfläche" an.
- Ein Menü öffnet sich, hier klicken Sie bitte auf „Change a password (=Kennwort ändern)" (siehe Abbildung 17).
- Anschließend öffnet sich eine Seite, auf der Sie Ihr Kennwort ändern können (siehe Abbildung 18). Tragen Sie bitte im ersten Feld Ihre Kennung ein (fdz Zahl Zahl Zahl Zahl Buchstabe Nachname Buchstabe Vorname weitere Zahlen, z.B. fdz9999dj001).
- Im zweiten Feld tragen Sie bitte Ihr altes Passwort und im dritten Feld Ihr neues Passwort ein, welches Sie dann im vierten Feld nochmals bestätigen.

Abbildung 18: Ändern des Passwortes: Passwort ändern



- Sie erhalten vom System anschließend eine Rückmeldung, dass Ihr Passwort geändert wurde. Schließen Sie den Vorgang mit "OK" ab. Sie werden wieder auf Ihren Desktop zurückgeleitet.
- Wiederholen Sie dieses Vorgehen ggf. für alle Ihre aktiven Projekte. Bitte beachten Sie: Sie müssen dazu nicht das Projekt wechseln, sondern Sie können von jedem Ihrer Projekte aus alle Passwörter Ihrer anderen Projekte ändern. Tragen Sie dazu den Namen des entsprechenden Benutzerkontos ein.

11.4.1 Passwortkonventionen

Ihr Passwort muss folgende Vorgaben erfüllen:

- Ihr Passwort muss mindestens 12 Zeichen lang sein.
- Ihr Passwort muss drei von vier Arten dieser Zeichen enthalten: Großbuchstaben, Kleinbuchstaben, Zahl, Sonderzeichen.

11.4.2 Kontosperre

- Bitte beachten Sie, dass sich Ihr Konto nach sechsmaliger Fehleingabe des Passworts sperrt.
- Sollten Sie sich gesperrt haben, geben Sie uns bitte Bescheid (iab.fdz@iab.de), wir entsperren Ihr Konto wieder.