

# Spezifikation Rohdatentransfer und Format zum Datenmanagementportal

---

## Transfer

---

Für den automatisierten dateibasierten Transfer von Messdaten in das Datenmanagementportal (DMP) des UFZ stehen verschiedene Endpunkte bereit. Transfers können via [Secure File Transfer Protocol \(SFTP\)](#), [File Transfer Protocol \(FTP\)](#) oder [Hypertext Transfer Protocol \(HTTP\)](#) übertragen werden. SFTP- und FTP-Transfers können aus dem Internet, HTTP-Transfers ausschließlich aus dem UFZ-Netzwerk initiiert werden. Bei der Übertragung über Mobilfunk wird die Nutzung spezieller Datenkarten (M-VPN-SIM) des UFZ empfohlen, mit denen eine direkte (eingeschränkte) Verbindung in das UFZ-Netzwerk ohne zusätzliche technische Voraussetzung wie VPN-Client oder -Router möglich ist. Generell muss der Transfer aktiv vom Datenlogger ausgehen. Es sind keinen Methoden zum zyklischen abholen der Messdaten vorgesehen.

Geräte mit direktem Netzwerkzugang sollten die Konfiguration per [DHCP](#) unterstützen. Mobilfunkmodems müssen [PPP](#) und die Konfiguration eines individuellen APN [6](#) unterstützen. Die Auflösung der IP-Adressen der verschiedenen Endpunkte muss dynamisch per [DNS](#) erfolgen.

## Secure File Transfer Protocol (SFTP)

Der zuverlässigste Weg die dateibasierten Rohdaten zu übertragen ist der Transfer über SFTP. Die Transportverschlüsselung sichert Authentizität, Vertraulichkeit und Integrität der Daten und schützt damit vor Missbrauch und Übertragungsfehlern. SFTP-Clients stehen für verschiedene Umgebungen und sämtliche relevanten Betriebssysteme zur Verfügung. Durch die Verschlüsselung entstehen erhöhte Anforderungen an die Rechenleistung des Clients, weswegen eine Implementierung nicht in jeder Umgebung möglich ist.

Weitere Informationen und Verweise auf einen Antrag zum Anlegen von SFTP-Accounts sind im UFZ-Intranet zu finden:

<https://www.intranet.ufz.de/index.php?de=33977>

## File Transfer Protocol (FTP)

Anders als beim Transfer über SFTP werden bei Übertragungen per FTP sowohl die Zugangs- als auch die Nutzdaten unverschlüsselt übertragen. FTP-Clients stehen für verschiedene Umgebungen und sämtliche relevanten Betriebssysteme zur Verfügung. Wegen der geringeren Anforderungen an die Rechenleistung ist eine Implementierung auch in sehr einfachen und auf sehr geringen Energieverbrauch optimierten Umgebungen möglich.

Weitere Informationen und Verweise auf einen Antrag zum Anlegen von FTP-Accounts sind im UFZ-Intranet zu finden:

<https://www.intranet.ufz.de/index.php?de=33977>

## Hypertext Transfer Protocol (HTTP)

Per HTTP können die Daten sowohl ver- als auch unverschlüsselt übertragen werden. Die Daten werden per HTTP-POST Anfrage an einen Server im UFZ-Netz ([logger-worker.intranet.ufz.de](http://logger-worker.intranet.ufz.de)) übertragen. Die

zu übertragende Datei wird dabei als Parameter mit dem Namen "file" eingebettet. Die Zuordnung zu einem bestimmten Logger erfolgt über eine Device-ID, die im DMP vergeben wird.

## Beispielimplementierungen

### Python

```
import requests

device_id = 'fffffffffffffff'
file      = 'test.csv'

paramter  = 'file'
base_url  = 'https://logger-worker.intranet.ufz.de/gateway/upload/device/'
url       = base_url + device_id

response = requests.post(url, files={
    paramter: open(file, 'rb')
})

ret      = response.json()
success  = response.status_code == requests.codes.ok

if len(ret['content']):
    print(ret['content'])

if not success and len(ret['exception']):
    print(ret['exception'])
    exit(1)
```

### Curl/Bash

```
#!/usr/bin/env bash

FILE="/tmp/test.csv"
DEVICE_ID="fffffffffffffff"

BASE_URL="https://logger-worker.intranet.ufz.de/gateway/upload/device/"
URL="${BASE_URL}${DEVICE_ID}"

curl "$URL" -sf -F "file=@${FILE}" > /dev/null
exit $?
```

### .NET/Powershell

```
$file      = "C:\tmp\test.csv"
$deviceid  = "fffffffffffffff"
```

```
$base_url = "https://logger-worker.intranet.ufz.de/gateway/upload/device/"

$url      = $base_url + $deviceid
$webClient = New-Object System.Net.WebClient;

$webClient.UploadFile($url, "POST", $file);
```

## Format

Die Rohdaten sollten UTF-8 kodiert in Form von CSV Text-Dateien übertragen werden.

## Technische Rahmenbedingungen

- Eine Datei sollte mindestens einen Zeitstempel enthalten mit mindestens einem Messwert enthalten
- Einzelne Dateien dürfen eine Länge von maximal 100.000 Zeilen nicht überschreiten
- Einzelne Dateien dürfen eine Größe von maximal 10 Megabyte nicht überschreiten
- Felder haben eine maximale Länge von 200 Zeichen
- Felder ohne Wert
  - sind entweder leer,
  - enthalten Leerzeichen
  - oder die Zeichenkette "NIL"
  - Beispiel:

```
timestamp, AirTemperature, Surfacewater, Battery
2013-09-12 12:23:33, 20.3,, 11.7
2013-09-12 12:23:33, 20.3, , 11.7
2013-09-12 12:23:33, 20.3, NIL, 11.7
```

Alle Zeilen enthalten in der dritten Spalte einen gültigen Nullwert

## Feldtrenner

Komma (",")

## Zeilentrenner

Konfigurierbar (siehe <http://de.wikipedia.org/wiki/Zeilenumbruch>)

```
Unix: Line feed ("LF", "0A" oder "\n")
Windows: Carriage return line feed ("CRLF", "0D0A" oder "\r\n")
```

## Dezimalpunkt

Punkt (".")

## Zeitstempel

In der ersten Spalte/Feld

Format "YYYY-MM-DD HH:mi:ss",  
zum Beispiel 2013-02-19 10:50:00

## Erste Zeile/Überschrift

Spaltenüberschriften (oder leer), wird ignoriert

## Zeichenketten

Zeichenketten dürfen keine Kommas enthalten.