

Art und Umfang der Leistung (Leistungsbeschreibung)

1. Gegenstand der Leistung

Das Projekt VirHand - Virtuelle Realität zur Unterstützung des Handwerks etabliert technologische, wirtschaftliche und soziale Innovationen als Schlüssel zum erfolgreichen Strukturwandel in der Lausitz. Ziel ist es, ein attraktives und an Nachhaltigkeitszielen ausgerichtetes Handwerk in der Region mitzugestalten, in dem lebenslanges Lernen erfolgreich praktiziert wird.

Gegenstand der Leistung ist die Bereitstellung eines strukturierten Datensatzes zu Stellenausschreibungen in Deutschland für den Zeitraum 01.01.2016 bis 31.12.2025. Der Datensatz wird für eine wissenschaftliche Studie verwendet und muss eine zeitliche und regionale Analyse des deutschen Arbeitsmarktes ermöglichen.

2. Zeitliche, räumliche und mengenmäßige Abdeckung

- Abgedeckter Zeitraum: 2016–2025
- Mindestens wöchentliche Aktualisierung
- Mindestens 1.000 Stellenausschreibungen pro Jahr
- Deutschlandweit
- Regionale Abgrenzung auf Basis von 5-stelligen Postleitzahlen oder vergleichbar

3. Grundstruktur des Datensatzes

Der Datensatz muss:

- geordnet und maschinenlesbar sein,
- für jede Stellenausschreibung eine eindeutige Anzeigen-ID enthalten,
- eine eindeutige Zuordnung zum ausschreibenden Unternehmen ermöglichen,
- Verzerrungen durch Tochterunternehmen vermeiden (z. B. konsistente Zuordnung auf Unternehmensebene).

4. Mindestinhalte je Stellenausschreibung (Pflichtfelder)

Für jede Stellenausschreibung sind mindestens folgende Informationen bereitzustellen:

4.1 Zeit und Identifikation

- Eindeutige Anzeigen-ID
- Veröffentlichungsort bzw. -plattform
- Veröffentlichungsdatum der Anzeige
- Veröffentlichungsdauer der Anzeige

4.2 Arbeitgeber

- Unternehmensname (rechtlicher oder gebräuchlicher Name)
- Eindeutige Unternehmenszuordnung (stabile Unternehmens-ID oder konsistente Namenslogik)

4.3 Standort

- Postleitzahl (5-stellig)
 - Ort, Bundesland
-

4.4 Stelleninformationen

- Stellentitel
 - Jobbranche
 - Anstellungsart (z.B. Vollzeit, Teilzeit, Festanstellung, befristeter Vertrag, Ausbildung, Praktikum, Werkstudent, Minijob)
 - Vollständiger Text der Stellenausschreibung (Rohtext)
-

5. Datenqualität und Aufbereitung

- Der Datensatz muss frei von offensichtlichen Dubletten sein oder Dubletten eindeutig kennzeichnen
 - Mehrfachveröffentlichungen derselben Stelle (z. B. über verschiedene Plattformen) sollen nach Möglichkeit erkennbar sein
 - Fehlende Informationen müssen klar als fehlend gekennzeichnet werden
 - Es erfolgt keine verdeckte inhaltliche Veränderung oder Kürzung der Anzeigentexte
-

6. Datenformat und Lieferung

- Bereitstellung in einem maschinenlesbaren Format (z. B. CSV oder vergleichbar)
 - Einheitliche Struktur über den gesamten Zeitraum inkl. kurzer Beschreibung
 - Lieferung eines vollständigen historischen Datensatzes
-

7. Nutzungsrechte

- Der Datensatz muss uneingeschränkt für wissenschaftliche Forschung nutzbar sein
 - Analyse, Weiterverarbeitung und Veröffentlichung aggregierter Ergebnisse müssen erlaubt sein
 - Es dürfen keine inhaltlichen oder zeitlichen Nutzungseinschränkungen bestehen
-

8. Ausschlusskriterien

Ein Angebot ist vom Vergabeverfahren auszuschließen, wenn mindestens eines der folgenden Kriterien nicht erfüllt ist:

- Der angebotene Datensatz deckt den Zeitraum 2016–2025 nicht vollständig ab.
- Der Datensatz enthält weniger als 500 Stellenausschreibungen bei mehr als einem Kalenderjahr.
- Eine regionale Zuordnung auf Basis von 5-stelligen Postleitzahlen (oder vergleichbar) ist nicht möglich.
- Für Stellenausschreibungen fehlt der vollständige Anzeigentext (Rohtext).
- Es ist keine eindeutige Anzeigen-ID je Stellenausschreibung vorhanden.
- Eine eineindeutige Zuordnung zu Unternehmen ist nicht möglich oder führt zu systematischen Verzerrungen durch Tochterunternehmen.
- Die uneingeschränkte Nutzung für wissenschaftliche Forschung wird nicht eingeräumt.