

# Factsheet

Tool	
<b>Tool-Name</b>	Transkribus Lite
<b>URI</b>	<a href="https://app.transkribus.eu/">https://app.transkribus.eu/</a>
<b>Letzter Zugriff</b>	22.06.2023
<b>Reviewer:in</b>	Georg Maximilian Reiter
<b>Scope</b>	Review als Tool-Nutzer

System	
<b>Scope des Tools</b>	Primär KI-gestützte Layout- und Texterkennung von gedruckten und handgeschriebenen Texten, Training von Layout- und Texterkennungsmodellen. Annotationen mit Einschränkungen möglich
<b>Softwareumgebung/Softwaretyp</b> (Remotesystem im Browser / Lokaler Client)	Browser-Anwendung/ web-basiert
<b>Unterstützte Plattformen</b>	Windows, Linux, Mac
<b>Geräte</b>	Desktop
<b>Einbindung anderer Systeme (Interoperabilität)</b>	✗ (Private FTP, IIIF Manifest und DFG Viewer METS derzeit nur bei Transkribus eXpert)
<b>Accountsystem</b>	✓
<b>Kostenmodell</b> (Kostenübersicht/Open Source)	Layouterkennung und Training von Layout- und Texterkennungsmodellen, Upload, manuelle Transkription, Anreicherung und Download der Daten kostenfrei  Creditsystem für

	<p>automatische Texterkennung (1 Credit = 1 handgeschriebene Seite oder 1 Credit = 6 gedruckte Seiten)</p> <p>500 Credits bei Registrierung kostenlos (= 500 handgeschriebene oder 3000 gedruckte Seiten)</p>
<b>Anforderungen &amp; Methoden</b>	
<b>Erforderte Code Literacy</b>	✗
<b>Interface-Sprachen (ISO 639-1)</b>	de, en, es, et, fi, fr, it, nl, pl, pt, sl, sv
<b>Unterstützte Zeichenkodierung</b>	Unicode (und einige speziellere historische Kodierungen, z. B. Medieval Unicode Font Initiative)
<b>Inkludierte Datenkonvertierung</b> (Im Preprocessing mögliche Anpassung der Daten an für die Software erforderliches Format )	✗
<b>Abhängigkeit von anderer Software</b> (Falls ja, wird diese Software automatisch mitinstalliert?)	✗
<b>Erforderliche Plug-Ins</b> (bei web-basierten Anwendungen)	✗
<b>Dokumentation und Support</b>	
<b>Wartung und ständige Erweiterung</b>	✓
<b>Einbindung der Community</b>	<p>✓ Privatpersonen und Institutionen können Mitglied der Genossenschaft READ-COOP SCE und damit Miteigentümer des Unternehmens werden (Vorteile: Rabatte, 500 freie Credits pro Jahr, Erhalt privilegierter Informationen, Stimmrecht u. a.)</p>
<b>Dokumentation</b>	✓
<b>Dokumentationssprache</b>	HelpCenter (Englisch, Deutsch, Italienisch), How-to-Guides (Englisch),

	Glossar (Englisch), FAQ (Englisch, Deutsch, Italienisch)
<b>Dokumentationsformat</b>	HTML
<b>Dokumentationsabschnitte</b>	Ausführliche Dokumentation im <a href="#">HelpCenter</a> (Erste Schritte, Texterkennung, Layouterkennung, Händisch transkribieren, Modelltraining, Tagging, Suchen, Verwalten, Herunterladen), <a href="#">How-to-Guides</a> , <a href="#">Glossar</a> , <a href="#">FAQ</a>
<b>Verfügbarkeit von Tutorials</b>	✓ Schritt-für-Schritt-Anleitungen (HelpCenter, How-to-Guides), instruktive mehrsprachige Einführungswebinare sowie weitere Videos auf <a href="#">YouTube</a> verfügbar
<b>Aktiver Support/Community</b> (Forum, Slack, Issue Tracker etc.)	✓ Kontaktformular, aktive, aber nicht von Transkribus administrierte <a href="#">Facebook-Gruppe</a>
<b>Nutzbarkeit &amp; Nachhaltigkeit</b>	
<b>Installationsablauf</b>	keine Installation nötig
<b>Test</b> (Gibt es ein Test Suite, um zu überprüfen, ob die Installation erfolgreich war?)	[nicht anwendbar]
<b>Lizenz, unter der das Tool veröffentlicht wurde</b>	Proprietäre Software
<b>Registrierung in einem Repository</b>	✗
<b>Möglichkeit, zur Software-Entwicklung beizutragen</b>	✗
<b>Benutzerinteraktion &amp; Benutzeroberfläche</b>	
<b>Benutzerprofil</b> (Erwartete Nutzer:innen)	Primär GeWi-Forschungsinstitutionen und Forschende als Tool-Nutzende
<b>Benutzerinteraktion</b> (erwartete Nutzung)	Anlegen von Sammlungen und Dokumenten, Hochladen von

	Dateien, KI-gestützte Layout- und Texterkennung, Training von Layout- und Texterkennungsmodellen, Anreicherung von Texten
<b>Benutzeroberfläche</b>	Webbasiertes GUI
<b>Visualisierungen</b> (Analyse-, Input-, Outputkonfigurationen)	✗
<b>Benutzerverwaltung</b>	
<b>Personenverwaltung</b>	✓ Hinzufügen/Einschränkung von Mitarbeitenden möglich
<b>Interne Kommunikationsmöglichkeiten</b> (z. B. Annotationsrichtlinien, Kommentarfunktionen ...)	✗
<b>Daten- und Toolverwaltung</b>	
<b>Zentrale/dezentrale Verwaltungsmöglichkeit</b>	✓ Mehrere Sammlungseigentümer:innen möglich
<b>Versionskontrolle</b>	✓ Bei den einzelnen transkribierten Seiten
<b>Projektspezifische Einstellungen</b> (eigene Metadaten, Tags anlegbar etc.)	✓
<b>API</b>	✓ MetagraphoAPI
<b>Möglichkeit des simultanen Arbeitens</b>	✓ Nicht simultan an einzelnen Seiten, aber Dokumenten und Sammlungen
<b>Datenupload</b>	
<b>Unterstützte Dateiformate</b>	Bilder (JPEG, PNG) oder PDFs
<b>Informationen zur Datensicherheit</b>	Alle hochgeladenen Dokumente sind standardmäßig privat. Die Daten werden auf Servern in Innsbruck (Österreich) in DSGVO-konformer Weise gespeichert und können gemäß den <u>Terms &amp;</u>

	<u>Conditions</u> der READ-COOP SCE verarbeitet werden
<b>Zugänglichkeit von verschiedenen Standorten/Geräten</b>	✓
<b>Einschränkungen hinsichtlich der Datenmenge</b>	✗
<b>Verlustfreier Upload von bereits bearbeiteten Dokumenten</b>	[nicht anwendbar, da noch keine Annotation stattgefunden hat]
<b>Unterstützung von IIIF-Import</b>	✗ (derzeit IIIF Manifest nur in Transkribus eXpert unterstützt)
<b>Datenbearbeitung (Transkriptionstool)</b>	
<b>Komplexitätsgrad beim Mark-Up</b> (z.B. Verfügbarkeit von Buttons, Drag&Drop-Funktion)	Strukturtags (Faksimilie) und Texttags, Formatauszeichnung (fett, kursiv usw.)
<b>Standardeinstellungen entsprechend bestimmten Standards für Digitale Editionen</b>	✗
<b>Anpassungsmöglichkeit und Validierung entsprechend projektspezifischen Konventionen/Schemata</b>	✓ Tags an TEI anpassbar, allerdings keine Validierungsmöglichkeit
<b>Definition eigener/projektspezifischer Tags</b>	✓
<b>Metadaten-Anreicherung</b>	✓ (Sammlungs- und Dokumentmetadaten)
<b>Layoutmöglichkeiten</b> (z.B. Tabellendarstellung, Spalten wie in Zeitschriften ...)	✗
<b>Eigene Indexierung</b>	✗
<b>Möglichkeit von Textvergleich bzw. Arbeit an Variantenapparat</b>	✗
<b>Ansichtsmöglichkeiten</b> (z. B. Bearbeitungsansicht, Vorschau, Synopsen-Ansicht ...)	Bild-Texteditor-Synopse
<b>Verlinkung von Entitäten, NER</b>	✗
<b>Datenexport</b>	
<b>Unterstützte Dateiformate</b>	PAGE XML, ALTO XML, TEI XML, PDF (Bild- und

	Transkriptionslayer), Docx, Tags XLSX, Table XLSX
<b>Datenverlust</b> (nicht vollständiger Erhalt von Annotationen, die bereits vor Verwendung des Tools gemacht wurden)	[nicht anwendbar]
<b>Validierungsmöglichkeit für TEI-XML vor Export</b>	✗
<b>Datenaufbewahrung nach Export</b>	Sammlungen, in denen 120 keine Aktivität stattfand, werden archiviert und an die Eigentümer:innen wird ein Mail versandt. Finden dann innerhalb der nächsten 21 Tag keine Aktivitäten statt, werden die Sammlungen gelöscht