

Factsheet

Tool	
Tool-Name	Transkribus Lite
URI	https://app.transkribus.eu/
Letzter Zugriff	22.06.2023
Reviewer:in	Georg Maximilian Reiter
Scope	Review als Tool-Nutzer

System	
Scope des Tools	Primär KI-gestützte Layout- und Texterkennung von gedruckten und handgeschriebenen Texten, Training von Layout- und Texterkennungsmodellen. Annotationen mit Einschränkungen möglich
Softwareumgebung/Softwaretyp (Remotesystem im Browser / Lokaler Client)	Browser-Anwendung/ web-basiert
Unterstützte Plattformen	Windows, Linux, Mac
Geräte	Desktop
Einbindung anderer Systeme (Interoperabilität)	✗ (Private FTP, IIIF Manifest und DFG Viewer METS derzeit nur bei Transkribus eXpert)
Accountsystem	✓
Kostenmodell (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>
Anforderungen & Methoden	
Erforderte Code Literacy	✗
Interface-Sprachen (ISO 639-1)	de, en, es, et, fi, fr, it, nl, pl, pt, sl, sv
Unterstützte Zeichenkodierung	UTF-8
Inkludierte Datenkonvertierung (Im Preprocessing mögliche Anpassung der Daten an für die Software erforderliches Format)	✗
Abhängigkeit von anderer Software (Falls ja, wird diese Software automatisch mitinstalliert?)	✗
Erforderliche Plug-Ins (bei web-basierten Anwendungen)	✗
Dokumentation und Support	
Wartung und ständige Erweiterung	✓
Einbindung der Community	✓ 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.)
Dokumentation	✓
Dokumentationssprache	HelpCenter (Englisch, Deutsch, Italienisch), How-to-Guides (Englisch), Glossar (Englisch), FAQ (Englisch, Deutsch, Italienisch)

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

	Anreicherung von Texten
Benutzeroberfläche	Webbasiertes GUI
Visualisierungen (Analyse-, Input-, Outputkonfigurationen)	✗
Benutzerverwaltung	
Personenverwaltung	✓ Hinzufügen/Einschränkung von Mitarbeitenden möglich
Interne Kommunikationsmöglichkeiten (z. B. Annotationsrichtlinien, Kommentarfunktionen ...)	✗
Daten- und Toolverwaltung	
Zentrale/dezentrale Verwaltungsmöglichkeit	✓ Mehrere Sammlungseigentümer:innen möglich
Versionskontrolle	✓ Bei den einzelnen transkribierten Seiten
Projektspezifische Einstellungen (eigene Metadaten, Tags anlegbar etc.)	✓
API	✓ MetagraphoAPI
Möglichkeit des simultanen Arbeitens	✓ Nicht simultan an einzelnen Seiten, aber Dokumenten und Sammlungen
Datenupload	
Unterstützte Dateiformate	Bilder (JPEG, PNG) oder PDFs
Informationen zur Datensicherheit	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 & Conditions</u> der READ-COOP SCE verarbeitet werden
Zugänglichkeit von verschiedenen Standorten/Geräten	✓

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

(nicht vollständiger Erhalt von Annotationen, die bereits vor Verwendung des Tools gemacht wurden)	
Validierungsmöglichkeit für TEI-XML vor Export	✗
Datenaufbewahrung nach Export	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