Nutzungsanleitung für eScriptorium

Inhaltsverzeichnis

1. Schritt fü	ir Schritt
1.1. Einl	oggen in eScriptorium
1.2. Ein	neues Projekt anlegen
1.2.1.	Terminologie
1.2.2.	Instruktionen
1.3. Ein	neues Dokument anlegen 4
1.3.1.	Terminologie4
1.3.2.	Instruktionen
1.4. Bild	er hochladen4
1.4.1.	Zugriff auf das Interface 4
1.4.2.	Lokale Dateien importieren5
1.4.3.	Bilder aus einer PDF Datei importieren5
1.4.4.	Verwendung eines IIIF-Manifests 6
1.5. Dok	umente manuell mit Annotationen versehen6
1.5.1.	Zugriff auf das Interface6
1.5.2.	Segmente und Bereiche auf dem Bild mit Annotationen versehen
1.5.3.	Annotieren der Transkription
1.5.4.	Eine Anmerkung zur Gliederung von Baselines, Polygonen und Transkriptionen9
1.5.5.	Die Reihenfolge der Zeilen ändern 10
1.5.6.	Semantische Annotationen 10
1.6. Ann	notationen Importieren 11
1.6.1.	Strukturierte Annotationen als XML importieren 11
1.6.2.	Importieren von Annotationen aus einfachem Text 11
1.7. Dok	umente automatisch mit Annotationen versehen12

	1.7.2	1.	Anleitung	12
	1.7.2	2.	Einrichten der Segmentierung	12
	1.7.3	3.	Vergleich verschiedener Transkriptionen	12
	1.8.	Mod	delle trainieren	13
	1.8.2	1.	Ein Training starten	13
	1.8.2	2.	Ein Modell verfeinern oder von Grund auf neu trainieren?	14
	1.9.	Ann	otationen exportieren	14
2.	Verv	valte	en einer kollaborativen Annotation Campaign	15
	2.1.	Erst	ellen einer Usergruppe (Admin)	15
	2.2.	Teile	en eines Dokumentes mit einem anderen User oder einer Gruppe	15
	2.3.	Eine	e Modell mit einem anderen User oder einer Gruppe teilen	15
3.	Sons	stiges	S	16

1. Schritt für Schritt

1.1. Einloggen in eScriptorium

Sie loggen sich bei eScriptorium mit einem individuellen Konto ein, das aus einem Login, einem Passwort und einer E-Mail-Adresse besteht. Ein Konto wird vom Benutzer nach Erhalt einer Einladung oder vom Administrator der Website erstellt:

S eScriptorium Home Kontakt	Anmelden
Bitte melden Sie sich an test-account	
•••••	
login	

Sobald Sie eingeloggt sind, sehen Sie das Dashboard, welches Ihnen alle Dokumente, die Sie selbst erstellt haben und die mit ihnen geteilt wurden, anzeigt. Beim ersten Login ist das Dashboard leer.

S eScriptorium Home Ko	Meine Projekte	Meine Modelle	
Meine Projekte		Erzeuge neu	ies Projekt

1.2. Ein neues Projekt anlegen

1.2.1. Terminologie

- Ein "**Projekt"** umfasst ein oder mehrere Dokumente, die Sie diesem Projekt zugeordnet haben.
 - 1.2.2. Instruktionen

Um ein neues Projekt zu erstellen, klicken Sie auf den Button "Erzeuge neues Projekt", dadurch öffnet sich eine neue Seite auf der Sie den Namen des Projektes angeben müssen. Anschließend erscheint Ihr neues Projekt in Ihrem Dashboard.

<mark>ട്ട</mark> eScriptorium Home Kontakt			Meine Projekte Meine Modelle	Hallo Iwill 👻
Meine Projekte			Project created successfully!	×
Test	lwill	0 Dokumente.		

1.3. Ein neues Dokument anlegen

1.3.1. Terminologie

- Ein "**Dokument**" ist eine Kollektion von Bildern, die eine Einheit bilden.
- Ein "Teil-Dokument" ist ein Bild bzw. eine Seite, die Teil eines Dokumentes ist.

1.3.2. Instruktionen

Um ein Dokument zu erstellen, klicken Sie innerhalb eines Projektes auf den Button "Neues Dokument anlegen", dadurch öffnet sich eine neue Seite sowie ein Eingabeformular:

Beschreibung Bilder Bearbeiten Modelle	
Test-Dokument	
Main script	
Latin	٠
Read direction	
Links nach rechts	÷
Die Lese-Richtung beschreibt die Anordnung der Elemente im Dokument. Im Unterschied dazu beschreibt die Text-Richtung die Anordnung der Wörter in ein	er Zeile und ist durch die Schriftart vorgegeben.
Line offset	
Baseline	٥
Die Position der Zeile in Abhängigkeit vom Polygon.	
Metadaten	Ontologie ?
Key Value	Die Standardkonfiguration enthält nur semantische Typen.
	Es wird nicht empfohlen, semantische, stilistische und Layout-Typen zu mischen.
	Sie müssen auf 'Aktualisieren' klicken, um Ihre neuen Typen dem Dokument hinzuzufügen.
	Bereichsstypen
	Commentary Illustration Main Title
	Einen Bereichstyp hinzufügen
	Zeilentypen
	Einen Zeilentyp hinzufügen
Erze	

Das Feld "Name" muss ausgefüllt werden. Alle anderen Felder sind optional und können später ergänzt werden.

Nach der Eingabe der Informationen, klicken Sie auf "Erzeugen", um das Dokument zu erstellen. Anschließend wird die Meldung "Document created successfully!" angezeigt. Die eingegebenen Daten können später immer noch angepasst werden: der "Erzeugen"-Button wird durch einen "Aktualisieren"-Button ersetzt.

SeSeriptorium Home Kontakt		Meine Projekte Meine Modelle Hallo Iwill 🔻
Demo Dokumente		Neues Dokument anlegen
Text_Dokument	Iwill 24. März 2022	13 Bilder

Um vom Dashboard wieder auf das Eingabeformular bzw. das zuletzt bearbeitete Element zu kommen, müssen Sie auf das "Edit"-Icon klicken, welches sich neben dem Dokument befindet oder auf den Namen.

- 1.4. Bilder hochladen
- 1.4.1. Zugriff auf das Interface

Unter der Schnittstelle "Bilder" werden alle Anwendungen hinsichtlich der automatischen Bildverarbeitung sowie der Importe und Exporte verwaltet.

Um auf diese Schnittstelle zuzugreifen, klicken Sie innerhalb eines Dokumentes einfach auf den Tab "Bilder".

- Beschreibung	g Bilder Bearb	eiten Modelle	Text_Dokument						
			1 Bilder hier ab	olegen oder zum Hochla	den anklicken.				
e auswählen 🛛 Alle	abwählen Ausgewäł	hlt 0/13	👌 import 😁 🕒 E	ixport	Trainieren 🔹	O B	inarisieren 🖹 Se	egment 🛛 📾 Tra	nskrib
1 × Tageback.	■ 2 ×	du Reise nuch Sie e ingsdoef, 1897	4 ×	5 ×	6 X		8 A	×	9

Es gibt mehrere Möglichkeiten Bilder auf die Plattform hochzuladen, die im Folgenden erläutert werden.

1.4.2. Lokale Dateien importieren

Die Bilder können einfach per "Drag and Drop" hochgeladen werden oder durch Auswählen der Bilder im File Explorer, mittels eines Klicks in die Box.

e	eScriptorium									Ξ	Ξ
_	÷	Beso	chreibung	Bilder B	earbeiten	Modelle	Text_Dok	ument			
	🇙 Bilder hier ablegen oder zum Hochladen anklicken.										
	All auswä	e ihlen	Alle abwählen	Ausgewählt 0/13	📲 Import 🔻	Export	🖽 Trainieren 🔻	O Binarisieren	≡ Segment	▲ Transkribieren	

Hinweis: Bevor die Seite neugeladen werden kann, muss der Import aller Bilder abgeschlossen sein. eScriptorium bietet momentan keine Möglichkeit Bilder automatisch zu sortieren, daher sollte beim Upload darauf geachtet werden, alle Bilder in der richtigen Reihenfolge auszuwählen sowie auf eine entsprechende Benennung zu achten ansonsten müssen die Bilder händisch sortiert werden.

1.4.3. Bilder aus einer PDF Datei importieren

Klicken Sie auf den "Import"-Button und dann auf "Images (PDF)" nun können sie eine PDF-Datei hochladen. Aus dieser werden automatisch Bilder extrahiert. Bitte beachten Sie, dass ausschließlich Bilder importiert werden. Sollte das PDF eine Textebene enthalten entsprechend der Transkription, wird diese nicht importiert.

1.4.4. Verwendung eines IIIF-Manifests

Klicken Sie auf den "Import"-Button, dann auf die Option "Images (IIIF)"und geben anschließend die URL des IIIF-Manifests ein. Alle Bilder werden lokal kopiert, so wie auch deren Metadaten (falls vorhanden), welche im Tab "Beschreibung" sichtbar sind. Beispiel-Link: <u>https://gallica.bnf.fr/iiif/ark:/12148/btv1b53026595r/manifest.json</u>

1.5. Dokumente manuell mit Annotationen versehen

1.5.1. Zugriff auf das Interface

Manuelle Annotationen sind nötig, um Ground-Truth-Daten zu generieren und damit Modelle zu trainieren oder um Ergebnisse dieser Operationen zu korrigieren. Es kann auch als Teil einer *Annotation Campaign* (gemeinsames Bearbeiten eines Dokumentes innerhalb einer Gruppe) verwendet werden, die nicht auf Kraken-Modelle zurückgreift (eScriptorium ist nur eine Input-Umgebung).

Unter dem Tab "Bearbeiten" können Sie Annotationen manuell erstellen und modifizieren. Diesen Button finden Sie am Seitenanfang sowie auf jedem einzelnen Bild in der Übersicht des Tabs "Bilder". Es gibt mehrere Bearbeitungsbereiche in diesem Tab passend zu den möglichen Editierungsoptionen (v.l.n.r.).

<mark>ős</mark> eSc	riptorium н	lome Konta	kt		Meine Projekte	Meine Modelle	Hallo Iwill 🔻
÷	Beschreibung	Bilder	Bearbeiten	Modelle	Text_Dokument Element 1 - UAT-407_105_1-15.pdf_page_1.png - (1991x2735) - 13.7 MB		
۹	>		Ì		A		
•							

- "Image Source" erlaubt es Ihnen das Quellbild anzusehen
- "Segmentation" zeigt den Segmentationsbearbeitungsbereich an
- **"Transcription"** zeigt den Bearbeitungsbereich der Transkription in der diplomatischen Ansicht an
- "Text Annotation" zeigt den Bearbeitungsbereich der Transkription im Textmodus an, um die Lesereihenfolge zu ändern oder um Transkriptionen zu erstellen

1.5.2. Segmente und Bereiche auf dem Bild mit Annotationen versehen

Im Segmentbearbeitungsfenster können Sie mehrere wesentliche Operationen durchführen:

Baselines:

- Das Zeichnen von "Baselines", die den Positionen des Textes auf dem Bild entsprechen, kann auf zwei unterschiedliche Arten erfolgen:
 - Freies Zeichnen (nicht empfohlen): linke Maustaste gedrückt halten und gewünschte Zeile ziehen

- Punkt-für-Punkt-Darstellung: Linksklick an die entsprechende Stelle, durch einen Rechtsklick können Punkte hinzugefügt werden und durch einen erneuten Linksklick die Zeile beendet werden
- Zeilen verschieben: Zeilen auswählen und mit STRG + Ziehen an der gewünschten Stelle platzieren
- Zeilen auswählen: SHIFT + Klick auf entsprechende Zeilen
- Auswahl aufheben: durch das Klicken auf eine leere Stelle, wird die Auswahl aufgehoben
- Zeilenzeichnen abbrechen: während die Zeile gezeichnet wird, Escape-Taste drücken
- Zeilen löschen: Der rote Papierkorb löscht alle ausgewählten Zeilen; wenn SHIFT gedrückt gehalten wird, können mehrere Zeilen gleichzeitig gelöscht werden
- Zeilen verbinden: Wenn mindestens zwei Zeilen ausgewählt sind, wird eine Option zum Verbinden der Zeilen verfügbar (J-Taste): SHIFT + Auswählen entsprechender Zeilen → SHIFT lösen → STRG + J
- Kontrollpunkte auf Zeile hinzufügen: Zeile auswählen und Doppel-Klick auf die entsprechende Stelle
- Verlängern mehrerer Zeilen: Falls mehrere Zeilen als zu kurz erkannt wurden, können Sie diese einfach auf eine Länge bringen (meist sinnvoll bei Texten im Blocksatz): Gelbe Schere aktivieren → SHIFT + Markieren aller Anfangs- bzw. Endpunkte der Reihe → Schere deaktivieren → SHIFT + erneut alle Punkte markieren → SHIFT lösen → STRG + alle Zeilen auf gewünschte Länge ziehen
- Kontrollpunkte auf einer Zeile auswählen oder verändern: Klicken Sie mit der linken Maustaste auf eine Zeile, um sie auszuwählen, und ziehen Sie dann den ihr am nächsten liegenden Kontrollpunkt. Ein Doppelklick auf die Zeile erzeugt einen neuen Kontrollpunkt an der Mausposition. Sie können die Reihenfolge der Punkte der Zeile umkehren, indem Sie eine Zeile (oder mehrere) auswählen und den Umkehrbutton (Taste I) betätigen.
- Verschieben einer oder mehrerer Kontrollpunkte: einfach Kontrollpunkt auswählen und an gewünschte Stelle ziehen
- Lasso-Auswahlwerkzeug: Mit SHIFT + Ziehen erzeugen Sie ein Lasso-Auswahlwerkzeug, womit Sie Kontrollpunkte auswählen können (sie erscheinen dann schwarz). Hinweis: Wenn bereits Zeilen ausgewählt sind, wählt das Lasso nur Punkte auf diesen Zeilen aus.
- Mehrere Kontrollpunkte bzw. Zeilen gleichzeitig verschieben: Mit der Tastenkombination STRG + Ziehen können Sie alle ausgewählten Kontrollpunkte auf einmal verschieben (oder die ausgewählten Zeilen, wenn keine Kontrollpunkte ausgewählt sind).
- Kontrollpunkte löschen: der gelbe Papierkorb löscht nur ausgewählte Kontrollpunkte; Hinweis: Es sollte eigentlich nur am Anfang und am Ende der Zeilen jeweils ein Punkt sein. Es könnte sein, dass bei einigen Dokumenten ein Problem bei der Transkription auftritt, falls zu viele Punkte vorhanden sind: SHIFT + innerhalb der Anfangs- und Endpunkte der Zeilen alles markieren → gelben Papierkorb anklicken.
- Scheren-Tool: Mit dem Ausschneidemodus (Taste C oder dem gelben Scheren-Button) können Sie Zeilen ausschneiden, in zwei Teile teilen oder einen Teil davon entfernen.
- Leserichtung einer Zeile: Die Leserichtung kann über den Button "Segment" im Tab "Bilder" festgelegt werden.

Masken:

• Berechnung von Polygonen/Masken: Aktivierung der Berechnung von Polygonen, die mit jeder Zeile verbunden sind (Vorgang ist automatisiert und wird asynchron verwaltet, ohne dass der Benutzer nach der ersten Verwendung etwas tun muss)

- Erstellung von neuen Masken: wenn eine Seite bereits Masken hat, erhalten neue Zeilen automatisch auch eine Maske. Bei der Aktualisierung einer Zeile, werden auch automatisch Masken neu berechnet.
- **Hinweis:** Die Qualität der Masken ist abhängig von der Qualität der Segmentierung und nicht nur der dazugehörigen Zeile. Es sollten alle Zeilen eingezeichnet sein, bevor die Masken kalkuliert werden.
- Nur Masken kalkulieren: Im Tab "Bilder" klicken Sie auf "Segment" und wählen dann die Option "Nur Zeilenmasken" aus.

Textbereiche:

In diesem Fenster können Bereiche (oder Zonen) erstellt werden und Segmente/Zeilen mit diesen verbunden werden. Ein Segment, das sich innerhalb eines Bereichs befindet, ist daher nicht automatisch mit diesem verbunden.

- Von Zeilenansicht in Bereichsansicht wechseln: "R"-Taste auf der Tastatur drücken
- Bereich erstellen: In der Bereichsansicht können Sie durch einen Linksklick einen neuen Bereich anlegen und durch einen erneuten Linksklick den ausgewählten Bereich anlegen/fertigstellen.
- **Bereiche verändern**: Bereiche können auf die gleiche Weise wie Kontrollpunkte ausgewählt und verändert werden.
- Zeilen, die innerhalb eines Bereichs gezeichnet werden, werden automatisch an diesen gebunden.
- Zeilen mit Bereichen verbinden bzw. davon trennen: Ausgewählte Zeilen können mit den entsprechenden Schaltflächen mit Bereichen verknüpft (Y-Taste) oder von ihnen getrennt (U-Taste) werden.
- Bereiche löschen: Durch anklicken des Bereichs in der Bereichsansicht und anschließendem anklicken des roten Papierkorbs, können Bereiche gelöscht werden.

Zusätzliche Hinweise:

Sie können jederzeit mit Strg + Z (Rückgängig machen) und Strg + Y (Wiederherstellen) Aktionen rückgängig machen bzw. wiederherstellen oder mit den entsprechenden Schaltflächen durch Ihren Änderungsverlauf gehen.

All diese Operationen, die in diesem Bereich genutzt werden können, finden Sie in einem Hilfsfenster, welches sich hinter dem ?-Button befindet.

1.5.3. Annotieren der Transkription

Nur wenn Baselines und Masken auf dem Bild festgelegt sind, gibt es die Möglichkeit die Funktionalitäten der Buttons "Transcription" and "Text" zu nutzen.

Um eine einer Zeile zugeordnete Transkription hinzuzufügen oder zu ändern, klicken Sie im Bereich "Transcription" auf die entsprechende Zeile. Ein Eingabefenster wird angezeigt. Um eine Transkription aufzunehmen, drücken Sie "Enter": es wird automatisch das Eingabefeld für die nächste Zeile angezeigt. Während Sie im Fenster "Transcription" tippen, werden die mit Zeilen versehenen Bereiche durch Text ersetzt und der Inhalt des Fensters "Text" ändert sich. Es ist also möglich den Text im Textfenster zu modifizieren, zu kopieren und mehrere Zeilen auf einmal einzufügen.



1.5.4. Eine Anmerkung zur Gliederung von Baselines, Polygonen und Transkriptionen

Die "Baseline" ist ein zentrales Element, um Informationen in der eScriptorium Database zu speichern. Also:

- Es ist möglich eine Baseline zu modifizieren (bewegen, Punkte hinzufügen) ohne, dass die Transkription beeinflusst wird
- Das Polygon wird immer von der Baseline aus berechnet, auch während des Trainings
- Es ist möglich das Polygon händisch zu ändern (nicht empfohlen), ohne dass dies Auswirkungen auf die Transkription hat



Während des Trainings der Kraken-Modelle, kann die Berechnung der Polygone zurückgesetzt werden: Für den User ist es daher von Vorteil, nicht in die Polygone einzugreifen und im Gegenteil dafür zu sorgen, dass die Baselines so gezeichnet werden, dass die automatisch erzeugten Polygone korrekt sind. Falls manuell Ground-Truth-Daten eingegeben werden, sollte darauf geachtet werden nur zu transkribieren was innerhalb des Polygons steht.

1.5.5. Die Reihenfolge der Zeilen ändern

Die Wiedergabereihenfolge der Zeilen erfolgt automatisch. Sie können sich die Ordnungsnummer jeder Zeile im Fenster "Segmentation" anzeigen lassen, indem sie auf "Toggle ordering display (L)" klicken, oder im Fenster "Text", wo die Zeilen in der folgenden Reihenfolge angezeigt werden.

Es ist möglich die Reihenfolge im "Text"-Fenster durch das Klicken von "Toggle sorting mode" zu ändern. Durch einfaches "Drag and Drop" der Zeilen kann die Änderung durchgeführt werden.

Hinweis: Es ist empfehlenswert, die Qualität der Segmentierung sicherzustellen, bevor die Reihenfolge der Zeilen geändert wird, weil das Hinzufügen und Entfernen von Zeilen die Berechnung dieser Reihenfolge systematisch neustartet und dabei manuelle Modifikationen überschreibt.



1.5.6. Semantische Annotationen

Es ist möglich, den Zeilen und Bereichen Etiketten (oder Tags) zuzuordnen, indem man einer vom User auf der Registerkarte "Beschreibung" vordefinierten Ontologie folgt. Es gibt Standard-Tags, aber es ist auch möglich welche über das Eingabefeld hinzuzufügen (Klicken Sie auf "+", dann "Aktualisieren" und anschließend fügen Sie den neuen Tag zur Liste hinzu) oder zu löschen (entfernen sie Die Haken von den Boxen vor dem Tag und klicken Sie anschließend auf "Aktualisieren").

Wählen Sie im Bereich "Segmentation" einen Bereich oder eine Zeile aus, klicken Sie auf "Set the type on all selected lines / regions (T)" und wählen Sie das entsprechende Tag aus. Die Farben des Bereichs oder Zeile ändern sich. Es ist möglich ein Tag auf mehrere Bereiche oder Zeilen auf einmal anzuwenden: dafür wählen Sie alle gewünschten Bereiche aus (STRG+ Klick und Ziehen oder STRG gedrückt halten und die gewünschten Zeilen/Bereiche anklicken).



1.6. Annotationen Importieren

1.6.1. Strukturierte Annotationen als XML importieren

Es gibt die Möglichkeit Segmentierungen oder Transkriptionen, die außerhalb von eScriptorium erstellt wurden zu importieren. Dafür klicken Sie unter dem Tab "Bilder" auf "Import" und wählen dann die Option "Transcription (XML)" aus.

Anschließend öffnet sich ein Formular und Sie können einen Namen für die importierte Version festlegen, und eine Datei für den Import hochladen. Dies kann eine ALTO XML Datei, eine PAGE XML Datei oder eine ZIP-Datei, die ALTO oder PAGE Dateien beinhaltet, sein.

Es ist nicht notwendig vorher auszuwählen, welche Teile des Dokuments vom Import betroffen sind: die Verbindung wird automatisch hergestellt durch die Informationen aus den XML Dateien.

	_
Siehe die Importdokumentation für Details.	×
Importieren Sie die Segmentierung und Transkriptionen aus XML-Datei.	
Name	
The name of the target transcription. Will default to '{format} Import'.	
□ Override existing segmentation.	
Destroys existing regions, lines and any bound transcription before impor	ting.
Durchsuchen Keine Datei ausgewählt.	
A single ALTO or PAGE XML file, or a zip file.	
Close Beginne II	mport
1 Er ingridert staged thighed and the	2 Balerra

Bitte beachten Sie, nachdem Segmentierungen importiert wurden, die nicht mit Kraken/eScriptorium erzeugt wurden, ist es wichtig die Polygone (Masken) zurückzusetzen bevor mit diesen Dokumenten Modelle trainiert werden.

1.6.2. Importieren von Annotationen aus einfachem Text

Es ist möglich manuell eine Transkription von einem einfachen Text zu importieren im Fenster "Text" innerhalb des Tabs "Bearbeiten". Jeder Zeilenumbruch zeigt dann den Übergang zu einem neuen Abschnitt an. Es ist daher darauf zu achten, dass die Lesereihenfolge der Zeilen übereinstimmt.

1.7. Dokumente automatisch mit Annotationen versehen

1.7.1. Anleitung

Automatische Dokumentannotationen werden über den Tab "Bilder" verwaltet.

- Wählen Sie die Bilder aus, die Sie mit Annotationen versehen wollen
- Klicken Sie auf "Segmentieren" (für die Erkennung von Baselines, Polygonen und/oder Bereichen) oder auf "Transkribieren" (für Transkriptionen)
- Ein Formular wird angezeigt: es erlaubt Ihnen ein Kraken-Modell hochzuladen oder ein Modell zu nutzen, das schon in eScriptorium existiert und für die Segmentierung, die Annotation zu konfigurieren

1.7.2. Einrichten der Segmentierung

- "Segmentation Steps" erlaubt es Ihnen festzulegen, auf welcher Ebene die Segmentierung durchgeführt werden soll
- "Zeilen und Bereiche" erzeugt Bereiche (sofern das verwendete Modell dafür trainiert wurde), Baselines und zugehörigen Polygonen
- "Zeilen Baselines und Masken" führt zur Erzeugung von Baselines und den damit verbundenen Polygonen
- Bei der Option **"Nur Zeilenmasken"** ist es nicht erforderlich, ein Modell zu laden; mit dieser Option können Sie die Berechnung der Polygone, die mit den bereits auf den Bildern vorhandenen Baselines verbunden sind, zurücksetzen.
- "Bereiche" ermöglichen es Baselines und Polygone, die schon auf den Bildern existieren zu erhalten und nur neue Bereiche zu generieren
- "Text direction" indiziert die Leserichtung der Zeilen

1 ausgewählte(s) Bild(er)	?
Achtung - eine Re-Segmentierung wird alle bestehenden Transkriptionen löschen!	×
Wählen Sie ein Modell aus	
Modell	
default (blla.mlmodel)	\$
Segmentation steps	
Zeilen und Bereiche	\$
Text direction	
Horizontal v. l. n. r.	\$
Override Wenn diese Option aktiviert ist, werden vorhandene Segmentierungen und verbundene Transkriptionen zuerst gelöscht!	damit
Close Segn	nent

- 1.7.3. Vergleich verschiedener Transkriptionen
- Laden Sie ein Dokument in eScriptorium hoch

- Führen Sie die Texterkennung mit dem gewünschten Modell durch und anschließend mit einem zweiten Modell oder mit manueller Texterkennung
- Nun wählen Sie die Transkription aus, die angezeigt werden soll
- Anschließend klicken Sie auf das Zahnrad-Symbol daneben und wählen die Transkription mit der Sie vergleichen möchten und schließen das Fenster wieder
- Nun schalten Sie die Transkriptionsansicht ein und klicken auf die entsprechende Zeile. Sie sehen nun markierten Text in rot und grün: Grüne Zeichen sind in der aktuellen (bearbeitbaren) Transkription nicht vorhanden, rote Zeichen sind in der verglichenen Transkription nicht vorhanden

G 🕤 🔳 Line #	<i>‡</i> 154	×			
Allgemeiner Inzeiger	tür Ladenburg und Schriesheim.				
Allgemeiner Anzeiger f by Iwill (eScriptorium) on Thu	Allgemeiner Anzeiger für Ladenburg und Schriesheim. by Iwill (eScriptorium) on Thu Feb 17 2022 12:16:19 GMT+0100				
-Toggle transcription c	omparison				
AEillgemeiner Anzeiger für JLadenburg und Schriesheim.	kraken:austriannewspapers_20220213164	3			
+Toggle history	1				

1.8. Modelle trainieren

1.8.1. Ein Training starten

Das Training von Kraken-Modellen startet man innerhalb des Tabs "Bilder". Es wird über den Tab "Modelle" verfolgt (inaktiv solange noch kein anderes Modell als das Standardmodell mit dem Dokument verknüpft ist).

Der Zugang zu den Trainingsfunktionen hängt von der Autorisierung ab. Dies kann über die Administrationsschnittstelle verwaltet werden.

Um ein Modell zu trainieren müssen sich alle Bilder/Annotationen im gleichen Dokument befinden.

- Wählen Sie die Dokumente aus, die die Ground-Truth-Daten enthalten
- Gehen Sie auf "Trainieren" und wählen Sie die Art des Trainingsmodells aus: "Segment" für ein Segmentierungsmodell, "Recognizer" für ein Transkriptionsmodell
- Füllen Sie das Formular aus und klicken Sie anschließend auf "Trainieren"
- Das Training wird gestartet, manchmal dauert es einen Augenblick warten bis dies angezeigt wird

• Wenn das Training beendet ist, erhalten Sie eine Benachrichtigung und können das finale Modell im Tab "Modelle" finden. Es ist nun möglich das Modell zu downloaden oder es auf einem Dokument anzuwenden.

1.8.2. Ein Modell verfeinern oder von Grund auf neu trainieren?

Das Formular für die Trainingskonfiguration ermöglicht Ihnen folgende Auswahl:

- **ob Sie von Grund auf neu beginnen möchten** (geben Sie hierfür einfach einen Namen für das zu erstellende Modell an)
- oder ob Sie ein Modell verfeinern möchten (es kann über den File Manager hochgeladen werden oder sich schon in der Modellliste des zugehörigen Dokuments befinden)

Achtung: In Version 0.6.9 ist das Bearbeiten der Namen von Modellen nur möglich, wenn man ein ganz neues Modell trainiert. Um zu vermeiden, dass ein bereits existierendes Modell überschrieben wird, sollten Sie das Modell lokal auf ihrem Rechner downloaden und umbenennen in den gewünschten Namen, dann können Sie es über das Formular erneut in eScriptorium hochladen.

1.9. Annotationen exportieren

Das Exportieren von Annotationen funktioniert über den Tab "Bilder".

- Wählen Sie die relevanten Bilder aus
- Klicken Sie auf "Export", dann füllen Sie das Formular aus:
 - Spezifizieren Sie die Version der Transkription, die Sie exportieren möchten
 - Spezifizieren Sie das Export-Format ("Alto" für XML ALTO, "Pagexml" für XML PAGE oder "Text" für einfachen Text)
 - Setzen Sie einen Haken bei "Include Images", wenn Sie zusätzlich die Bilder exportieren möchten
 - Klicken Sie auf "Export" und speichern sie die generierte ZIP-Datei

1 ausgewählte(s) Bild(er)	?			
(PLACEHOLDER)	×			
Transcription				
manual	\$			
File format				
ALTO	\$			
Include images Will significantly increase the time to produce and download the export.				
Region types				
✓ (Undefined region type) ✓ (Orphan lines)				
Close	port			

2. Verwalten einer kollaborativen Annotation Campaign

2.1. Erstellen einer Usergruppe (Admin)

Es ist möglich eine Nutzergruppe im Administrator-Dashboard (Vorausgesetzt Sie haben die erforderlichen Rechte) zu erstellen. Diese Gruppen dienen dazu Arbeitsgruppen zu definieren oder um ausgewählten Usern bestimmte Rechte zu erteilen.

Auch unter "Profil" \rightarrow "Teams" können Gruppen erstellt werden. Zudem werden hier auch die eigenen Gruppenzugehörigkeiten aufgelistet.

Informationen	Lege ein neues Team an		
API-Schlüssel	Name		Erzeugen
Dateien	Meine Teams		
Teams			
Einladungen	OCK-BM	Besitzer: admin	Verlassen

2.2. Teilen eines Dokumentes mit einem anderen User oder einer Gruppe

Ein User kann ein Dokument mit mehreren anderen Usern teilen, auch mit denjenigen, die nicht Teil der Gruppe sind, welcher das Dokument angehört. Dies geht über den Tab "Beschreibung":

- Klicken Sie auf den Button "Teile dieses Dokument"
- Geben Sie den Namen des Users ein, mit dem Sie Ihr Dokument teilen möchten oder setzen Sie einen Haken bei seinem Namen in der Liste
- Um zu bestätigen, klicken Sie auf "Teilen"

S eScriptorium Home Kontakt	Meine Projekte Meine Modelle Hallo Iwill 🔻
Content Beachreibung Bilder Bearbeiten Modelle	
Text_Dokument	
Main script	
Latin	\$
Read direction	
Links nach rechts	\$
Die Lese-Richtung beschreibt die Anordnung der Elemente im Dokument. Im Unterschied ist durch die Schriftart vorgegeben.	dazu beschreibt die Text-Richtung die Anordnung der Wörter in einer Zeile und
Line offset	
Baseline	\$
Die Position der Zeile in Abhängigkeit vom Polygon	

2.3. Eine Modell mit einem anderen User oder einer Gruppe teilen

Ein Modell ist mit einem Dokument verknüpft und nicht mit einem User. Um ein Modell mit anderen Usern zu teilen, gibt es zwei Möglichkeiten:

- Downloaden Sie das Modell im Tab "Modelle" und versenden Sie es über einen anderen Kanal (z.B. per E-Mail)
- Teilen Sie mit dem User das Dokument, dem das Modell zugeordnet ist. Nun kann der andere User es herunterladen und es in das Dokument laden, in dem er es anwenden möchte.

3. Sonstiges

- Wichtig: Zeilen, Bereiche und Masken immer bearbeiten bevor Transkription erfolgt ist, da diese sonst an dieser Stelle gelöscht werden können.
- **Binarisierung:** Dies ist meist nicht nötig, nach dem Hochladen der Bilder können diese i.d.R. direkt segmentiert werden. In den meisten Fällen verschlechtert die Binarisierung das Ergebnis.
- Keine Reaktion mehr bei der Bearbeitung des Dokuments: Seite erneut laden (tauchte bei Firefox bisher öfters auf, ist aber noch unklar wann genau es vorkommt und woran es liegt)
- Kein Warten im Fenster auf Segmentierung, Binarisierung und Transkription: Während dieser Prozesse, die mitunter länger dauern können, kann das Fenster verlassen werden. Der Prozess wird nicht abgebrochen.
- Kein Speichern nötig: Alle Vorgänge, werden automatisch gespeichert. Wird das Bearbeitungsfenster einmal verlassen, kann man Aktionen nicht mehr rückgängig machen.
- Sollte es Probleme mit der Internetverbindung geben, kann dies zum Verlust von Arbeitsschritten führen.