

Revrit

Automatisierte Retransliteration von Hebraica-Titeldaten

Auditing-Webtool

[Revrit](#) ermöglicht es, transliterierte Titeldaten von Hebraica mit Originalschrift anzureichern und verbessert Nachweis und Auffindbarkeit der betreffenden Werke. Das Verfahren wurde von dem an der Universitätsbibliothek Johann Christian Senckenberg angesiedelten [Fachinformationsdienst Jüdische Studien](#) entwickelt.

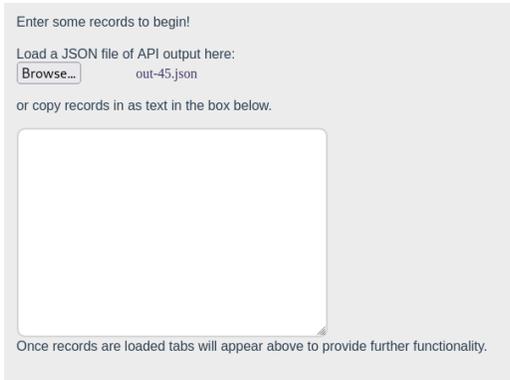
Revrit umfasst verschiedene Komponenten: eine JSON-API, Module zur Analyse, Verifikation & Diagnostik der Daten, das Konversionstool *Deromanize* sowie ein [Auditing-Webtool](#). Im Folgenden wird die Funktionalität des webbasierten Auditing-Tools beschrieben, das eine Begutachtung der Qualität der Konversionsergebnisse erlaubt sowie eine manuelle Korrektur von Titelkonversionen. Dies ist besonders relevant für nichtverifizierte Titeldatensätze. Grundsätzlich kann mit dem Auditing-Webtool jeder Datensatz bearbeitet werden.

Datenspeicherung und Datenschutz.....	1
1. Erste Schritte.....	2
2. Dateneingabe – <i>input records</i>	2
3. Selektion der zu bearbeitenden Daten – <i>filter input</i>	2
3.1. Selektion gemäß dem Konversionsergebnis	3
3.2. Selektion der wahrscheinlichen hebräischen Titel.....	3
3.3. Selektion gemäß den Empfehlungen für Konversionen des Typus <i>unverified</i>	3
4. Prüfung und Korrektur einzelner Datensätze – <i>audit individual records</i>	3
4.1. Anzeige des konvertierten Titels.....	3
4.2. Suchoptionen zur Überprüfung des Konversionsergebnisses.....	4
4.3. Link in die Referenzdatenbank.....	4
4.4. Bearbeitungsoptionen	4
4.5. Prüfung der verschiedenen Output-Kategorien.....	4
4.6. Modifikationen der Konversionsergebnisse – <i>edit title</i>	5
4.6.1. Standardeditor	5
4.6.2. einfacher Editor	5
5. Export der überprüften und bearbeiteten Daten – <i>export audited</i>	6
Links	6

Datenspeicherung und Datenschutz

Alle Daten, die mit dem Revrit-Auditing-Webtool bearbeitet werden, werden lokal im Browser der Anwender*in gespeichert und bei jedem späteren Aufruf des Tools aus diesem lokalen Speicher mit Hilfe der [Web Storage API](#) geladen. Eine Datenübertragung auf Server der Universitätsbibliothek JCS Frankfurt am Main findet nicht statt.

1. Erste Schritte



Wenn Sie das Revrit-Auditing-Webtool das erste Mal aufrufen, öffnet sich ein Eingabe-Bildschirm. Mit dem Upload von Datensätzen werden zusätzliche Funktionen angezeigt.

2. Dateneingabe – *input records*

Sie haben zwei Möglichkeiten, Datensätze ins Revrit-Auditing-Webtool einzuspielen:

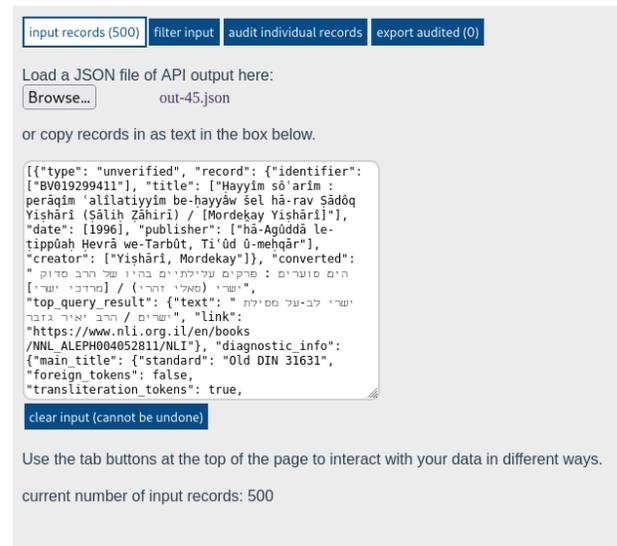
1. Datei-Upload via **Browse** bzw. **Durchsuchen**:

Eine JSON-Datei mit Revrit-Output kann je nach Rechnerkapazitäten mehrere tausend Datensätze umfassen. Wir empfehlen eine Begrenzung auf 500 Datensätze. Auf Wunsch liefert der FID Jüdische Studien partitionierte Dateien.

2. Eingabe der Datensätze durch manuelles Kopieren der entsprechenden Zeilen aus der JSON-Datei des Revrit-Outputs:

Die manuelle Eingabe erlaubt die Auswahl beliebiger Datensätze für das Auditing. Bitte achten Sie darauf, nur vollständige Zeilen einzuschließen, da fragmentierte Datensätze Fehler verursachen.

Die Schaltfläche **clear input** befindet sich am unteren Ende des Eingabebildschirms und erlaubt das Löschen der Eingabe. Diese Aktion kann nicht widerrufen werden.

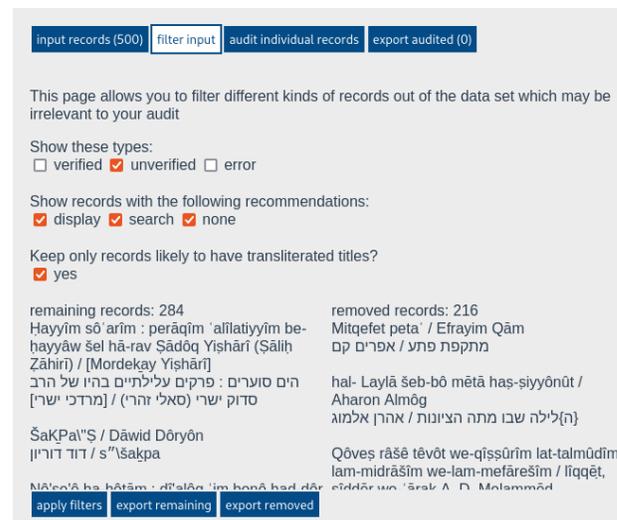


3. Selektion der zu bearbeitenden Daten – *filter input*

Über **filter input** können Sie Datensätze nach Outputkategorien (verified/unverified/error) und Empfehlungen (display/search/none) auswählen.

Bitte beachten Sie, dass die nicht selektierten Datensätze aus dem temporären Speicher des Revrit-Auditing-Webtools gelöscht werden und somit für die weitere Bearbeitung nicht mehr zur Verfügung stehen. Es besteht jedoch die Möglichkeit, die Filterergebnisse für spätere Bearbeitung zu exportieren und als separate Dateien zu speichern:

- **export remaining** – Export der selektierten Datensätze
- **export removed** – Export der nicht selektierten Datensätze



3.1. Selektion gemäß dem Konversionsergebnis

Standardmäßig sind Titelkonversionen des Typus *unverified* selektiert. Eine Überprüfung *verifizierter* Datensätze ist nicht zwingend notwendig. Titelkonversionen des Typus *error* werden als ungeeignet eingestuft. Beide Kategorien können dennoch manuell ausgewählt und geprüft bzw. bearbeitet werden.

Show these types:
 verified *unverified* error

3.2. Selektion der wahrscheinlichen hebräischen Titel

Die Revrit-API und das Revrit-Auditing-Webtool untersuchen die Zeichen des Eingabe-Titelstrings und beurteilen die Wahrscheinlichkeit, ob es sich um Transliterationen des Hebräischen handelt oder nicht. Standardmäßig werden nur Titel ausgewählt, die als wahrscheinliche Transliterationen des Hebräischen beurteilt werden.

Keep only records likely to have transliterated titles?
 yes

Hinweis: Diese Identifikation einer vorhandenen Transliteration ist nicht fehlerfrei, transliterierte Titel können fälschlicherweise aussortiert werden.

3.3. Selektion gemäß den Empfehlungen für Konversionen des Typus *unverified*

Die Qualität von *unverified* Titelkonversionen, für welche ein automatischer Abgleich mit der Referenzdatenbank nicht möglich war, ist unterschiedlich gut. Auf Basis der Analyse und Diagnostik der Daten gibt das Revrit-API Empfehlungen zur weiteren Prozessierung der Titelkonversion auf Ebene der Subfelder aus.

- *display* – Empfehlung zum Einlesen der Originalschrift für die Anzeige im Katalog
- *search* – Empfehlung zum Einlesen der Originalschrift in eine indizierte Titelkategorie zur Verbesserung der Katalogsuche
- *none* – keine Empfehlung zum automatischen Einlesen der Konversion. Derart gekennzeichnete Datensätze sollten unbedingt vor einem Einspielen in den Katalog mit dem Revrit-Auditing-Webtool überprüft und korrigiert werden.

Show records with the following recommendations:
 display search none

Unterhalb der Filteroptionen werden auf der linken Seite die Titel angezeigt, die für die Bearbeitung ausgewählt sind, und rechts diejenigen, die bei Filteranwendung entfernt werden.

Bitte beachten Sie, dass das Filtern durch Klick auf **apply filters** nicht rückgängig gemacht werden kann. Es empfiehlt sich, die zu entfernenden Daten zunächst zu exportieren, damit sie bei Bedarf bearbeitet werden können.

4. Prüfung und Korrektur einzelner Datensätze – *audit individual records*

4.1. Anzeige des konvertierten Titels



Über ***audit individual records*** kann das Bearbeitungs­menü aufgerufen werden. Es bietet zahlreiche weitere Funktionen, die mit der Anzeige des ersten Titels sichtbar werden.

4.2. Suchoptionen zur Überprüfung des Konversionsergebnisses

[search NLI](#) | [search KVK](#) | [search Google](#) | [Google Translate](#)

Unterhalb der angezeigten Titeldaten sind mehrere Suchoptionen für den konvertierten Titelstring verfügbar. Dies umfasst den Katalog der National Library of Israel (NLI), den Karlsruher Virtuellen Katalog sowie Google und Google Translate.

4.3. Link in die Referenzdatenbank

Unterhalb der Suchoptionen werden die Ergebnisse der Referenzdatenbank angezeigt:

- bei verifizierten Konversionen handelt es sich um den *matched title*
- bei nichtverifizierten Titeldaten wird das *top query result* angezeigt (siehe 4.1.)

matched title:
מתקפת פתע / אפרים קם

Der jeweilige Titel ist mit der Referenzdatenbank verlinkt. Derzeit handelt sich hierbei um den Katalog der NLI.

4.4. Bearbeitungsoptionen

[accept displayed title](#) [reject](#) [edit title](#) [show full record](#) [undo last](#)

- **accept displayed title** markiert das Konversionsergebnis zur Übernahme in den Zielkatalog
- **reject** markiert das Konversionsergebnis als inakzeptabel für die Übernahme in den Zielkatalog
- **edit title** erlaubt Modifikationen des Konversionsergebnisses
- **show full record** öffnet die Anzeige aller Ausgabedetails aus der Revrit-API
- **undo last** ermöglicht den Nutzenden bereits bearbeitete Titel erneut aufzurufen

4.5. Prüfung der verschiedenen Output-Kategorien

verified

Im Falle von verifizierten Datensätze konnte die originalschriftliche Konversion mit der Referenzdatenbank erfolgreich abgeglichen werden.

[input records \(499\)](#) [filter input](#) [audit individual records](#) [export audited \(2\)](#)

verified
Mitqefet peta^ʿ / Efrayim Qām

מתקפת פתע / אפרים קם

1990
Miśrad hab-Bittāhōn
Qām, Efrayim
[search NLI](#) | [search KVK](#) | [search Google](#) | [Google Translate](#)

matched title:
מתקפת פתע / אפרים קם

[accept displayed title](#) [reject](#) [edit title](#) [show full record](#) [undo last](#)

unverified

Wenn das Konversionsergebnis nicht mit der Referenzdatenbank abgeglichen werden konnte, wird ein Datensatz als *unverified* eingestuft. Das Auditing-Tool zeigt hier das *top query result* und verlinkt in die Referenzdatenbank.

[input records \(237\)](#) [filter input](#) [audit individual records](#) [export audited \(5\)](#)

unverified
Sefer Porat Yosef : ḥeleḳ 1 / Yosef Samiga

ספר פורת יוסף : חלק 1 / יוסף סמיגא

1590
Samigah, Yōsēf Ben-Binyāmīn
[search NLI](#) | [search KVK](#) | [search Google](#) | [Google Translate](#)

top query result:
... <<ספר>> פורת יוסף / שחיבר הגאון ... יוסף סמיגא על מסכת י"ט ...

[accept top query result](#)

[accept displayed title](#) [reject](#) [edit title](#) [show full record](#) [undo last](#)

error

Die Titel, die nichtlateinische Schriften enthalten oder andere unzulässige Elemente, führen zu einem Fehler. Das Auditing ermöglicht Prüfung und bei Bedarf Bearbeitung dieser Daten.



4.6. Modifikationen der Konversionsergebnisse – *edit title*

4.6.1. Standardeditor

Über **edit title** öffnen Sie das Bearbeitungsinterface, das über ein mouse-over für jedes einzelne Wort die alternativen Konversionsmöglichkeiten auflistet.



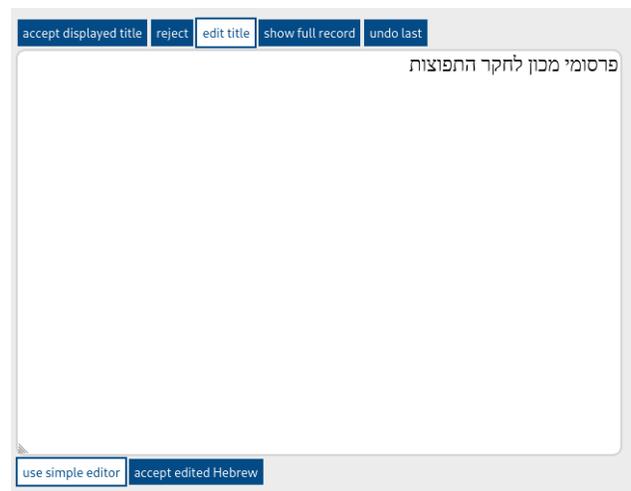
Sofern keine der alternativen Konversionsmöglichkeiten der korrekten Form entspricht, kann das betreffende Wort manuell bearbeitet werden. Ein Link zu Morfix erlaubt eine Überprüfung des jeweiligen Wortes.

Nach der Bearbeitung des Titels validieren Sie das Ergebnis über **accept edited Hebrew**. Der nächste Titel wird automatisch aufgerufen.

4.6.2. einfacher Editor

Eine Alternative zum Standardeditor ist der einfache Editor, der über **use simple editor** aufgerufen werden kann.

In der Textbox können Sie freien Text eintippen. Dieses Feature ist auch verfügbar, falls aufgrund technischer Probleme der Standardeditor nicht wie gewünscht reagieren sollte.



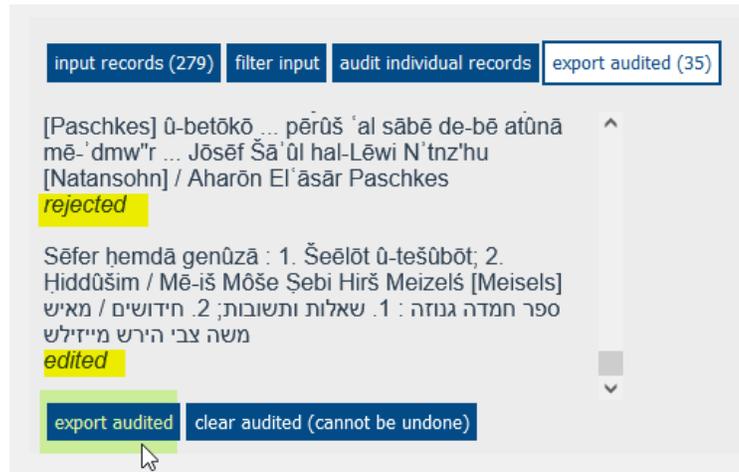
5. Export der überprüften und bearbeiteten Daten – *export audited*

Nach Prüfung und Bearbeitung der gewünschten Datensätze, können Sie die Ergebnisse über **export audited** exportieren.

Die Liste der untersuchten Titel enthält Transliteration, Rekonversion und Angabe zur Bearbeitung und Validierung.

Am Ende des Bearbeitungsmenüs ermöglicht **export audited** den Export der bearbeiteten Datensätze in eine neue JSON-Datei. Das Format der Datei entspricht dem Output der Revrit-API, enthält aber ein zusätzliches Feld **audit_result**, das sich in zwei Unterfelder aufteilt:

- **type** beschreibt die Art der Validierung des geprüften Titels
- **title** ist das finale Ergebnis und zeigt den bestätigten hebräischschriftlichen Titel an



The screenshot shows a web interface with a top navigation bar containing buttons for 'input records (279)', 'filter input', 'audit individual records', and 'export audited (35)'. Below this, a list of records is displayed. The first record is marked as 'rejected' and contains the text: '[Paschkes] û-betökô ... pêrûš 'al sâbē de-bē atûnâ mē-'dmw"r ... Jōsēf Šā'ûl hal-Lēwi N'tnz'hu [Natansohn] / Aharōn El'āsār Paschkes'. The second record is marked as 'edited' and contains: 'Sēfer hēmdā genûzā : 1. Šeēlōt û-tešûbōt; 2. Hiddûšim / Mē-iš Mōše Šebi Hirš Meizelš [Meizels] מֵאִשׁ / חִידוּשִׁים / 2. שְׁאֵלוֹת וְתַשׁוּבוֹת; 1. סֵפֶר חֲמֵדָה גְנוּזָה : 1. מִשָּׁה צְבִי הִירֵשׁ מִיִּזְיֵלֶשׁ'. At the bottom of the interface, there are two buttons: 'export audited' and 'clear audited (cannot be undone)'. A mouse cursor is pointing at the 'export audited' button.

Nach dem Export der Ergebnisse sollten Sie durch Klick auf **clear audited** den Speicher wieder freigeben. Dies ist insbesondere erforderlich, wenn die Bearbeitung mit weiteren Daten fortgesetzt werden soll. Diese Aktion kann nicht widerrufen werden kann.

Links

- | | |
|------------------------------|---|
| Link Revrit-Auditing-Webtool | https://api.jewishstudies.de/audit |
| Github-Dokumentation | https://github.com/ubffm/arcapi/blob/master/gui3/README.rst |
| Information zu Revrit | https://www.jewishstudies.de/revrit.html |
| Link Revrit-API | https://api.jewishstudies.de/ |

Kontakt

FID Jüdische Studien

Goethe-Universität Frankfurt am Main
 Bockenheimer Landstraße 134-138
 60325 Frankfurt am Main
 E-Mail: info@jewishstudies.de
 Web: www.jewishstudies.de

Projektleitung:

Dr. Kerstin von der Krone
 Tel.: 069/798-39665
 E-Mail: K.vonderKrone@ub.uni-frankfurt.de

Entwicklung

Aaron Christianson
 E-Mail: a.christianson@ub.uni-frankfurt.de

April 2023