

Medienrecherche v1.0 (Werkzeuge · Syntagma Forschung)

Serverloses Browser-Werkzeug für Recherche, Bearbeitung und Nachweis freier Medien

Strang: FORSCHUNG — Werkzeuge · Syntagma Forschung · Wien **Autor:** Thomas Schüller · ORCID 0009-0003-9799-6747 · forschung@tschueller.com **Lizenz:** CC BY 4.0 (Code und Dokumentation) **Stand:** August 2026 · **Version:** 1.0

Verifikationsvorbehalt. Die Angaben zu Schnittstellen, Zugriffsregeln (CORS), Schlüsselpflicht und Lizenzstatus der eingebundenen Archive beruhen auf einer systematischen Erhebung im Juli 2026. Sie sind Momentaufnahmen: Betreiber ändern Zugriffsregeln, Schlüsselpflichten und Nutzungsbedingungen ohne Vorlauf. Vor produktivem Einsatz und vor jeder Nachnutzung von Material ist der Status an der Primärquelle zu prüfen. Die Lizenz-Ampel des Werkzeugs bildet Metadaten der Archive ab und ersetzt keine rechtliche Prüfung.

1. Gegenstand

Medienrecherche ist ein Bündel von vier HTML-Dateien, die ohne Server, ohne Installation und ohne Übertragung von Inhalten im Browser laufen. Das Werkzeug verbindet drei Arbeitsschritte, die sonst auf getrennte Anwendungen verteilt sind:

1. **Suche** über mehrere offene Archive gleichzeitig, mit Filterung nach Medientyp und Lizenzstufe.
2. **Bearbeitung** gefundener Audiodateien (Schnitt, Entzerrung, Tonhöhe, Export) sowie Transkription zu MIDI und Notensatz.
3. **Nachweis** in Form einer exportierbaren Recherchedokumentation mit Lizenzbilanz.

Adressiert sind zwei Nutzungsformen gleichermaßen: journalistische Recherche und Medienarbeit einerseits, künstlerische Produktion andererseits. Beide benötigen dieselbe Grundlage — auffindbares Material mit geklärtem Nutzungsstatus und belegbarer Herkunft.

2. Einordnung

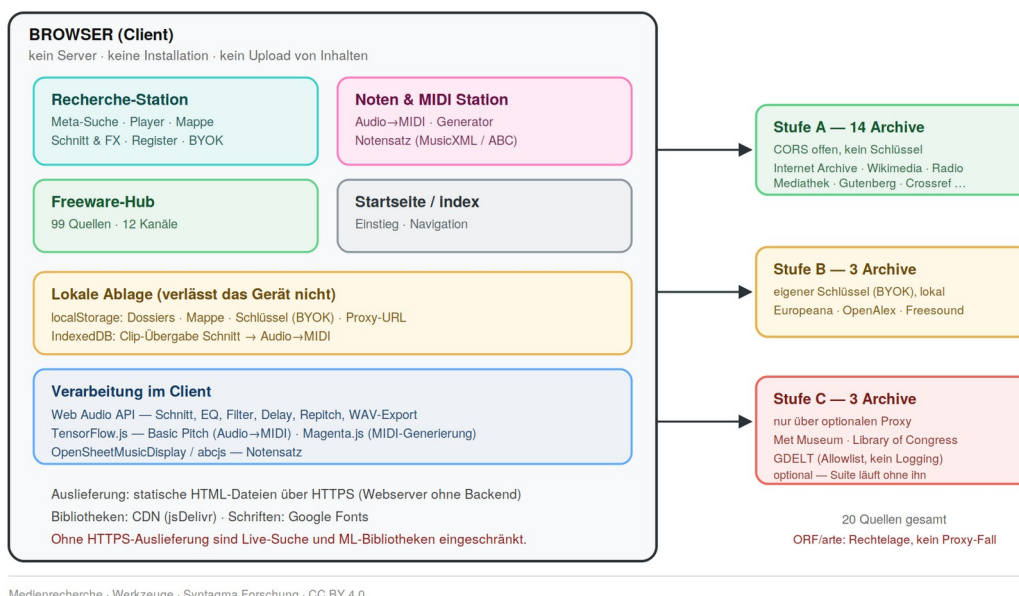
Bestehende Werkzeuge decken jeweils Teilbereiche ab: Suchmaschinen für Bilder mit Lizenzfilter, Sample-Bibliotheken, Notensatzprogramme, Verifikationswerkzeuge für Bildmaterial. Der Übergang zwischen diesen Schritten erfolgt bislang manuell, und der Lizenzstatus geht dabei regelmäßig verloren, weil er nicht mitgeführt wird.

Medienrecherche setzt an dieser Stelle an: Der Lizenzstatus wird vom Treffer bis in den Export durchgereicht, und die Übergabe zwischen Suche, Schnitt und Transkription erfolgt innerhalb derselben Anwendung. Die Entscheidung, welche Archive überhaupt eingebunden werden

können, folgt einem technischen Kriterium — der Frage, ob eine Schnittstelle aus statischem Browser-Code erreichbar ist.

Abb. 1 — Serverlose Architektur / Serverless architecture

Medienrecherche v1.0 · Ausführung vollständig im Browser / executed entirely in the browser



3. Architektur

3.1 Ausführungsmodell

Die Anwendung besteht aus statischen Dateien. Es gibt keine Datenhaltung auf einem Server, keine Benutzerkonten und keine Übertragung von Mediendateien. Verarbeitung geschieht im Browser über die Web Audio API und über Bibliotheken, die als JavaScript im Client laufen.

Lokal gespeichert werden ausschließlich Arbeitsdaten des Nutzers:

Ablage	Inhalt	Reichweite
localStorage	Dossiers, Sammelmappe, eigene Schlüssel, Proxy-Adresse	dieser Browser
IndexedDB	Zwischenablage für die Clip-Übergabe an die Transkription	dieser Browser

Eine Synchronisierung zwischen Geräten findet nicht statt. Dossiers und Mappe lassen sich als Datei exportieren und andernorts einlesen.

3.2 Auswahlkriterium der Quellen

Maßgeblich für die Einbindung ist nicht der Preis einer Schnittstelle, sondern ihre Erreichbarkeit aus dem Browser. Entscheidend sind zwei Eigenschaften: ob der Betreiber

Cross-Origin-Zugriffe erlaubt und ob ein Schlüssel erforderlich ist. Daraus ergeben sich drei Stufen.

Stufe A — unmittelbar nutzbar (14 Archive). Cross-Origin erlaubt, kein Schlüssel. Diese Archive stehen ohne jede Einrichtung zur Verfügung.

Archiv	Region	Medientyp
Internet Archive	global	Audio, Video, Text, Bild
Wikimedia Commons	global	Bild, Audio, Video
Art Institute Chicago	US	Bild (CC0)
Cleveland Museum of Art	US	Bild (CC0)
MediathekViewWeb (ARD, ZDF, ORF, arte u. a.)	DACH	Video
radio-browser	global	Audio (Streams)
Project Gutenberg	global	Text
Wikiquote	global	Text (Zitatnachweis)
LibriVox	global	Audio (Hörbuch, Bühnentext)
Chronicling America	US	Text (historische Presse)
Open Library	global	Text
Openverse	global	Bild, Audio
Crossref	global	Publikationen
Europe PMC	global	Publikationen

Stufe B — mit eigenem Schlüssel (3 Archive). Europeana, OpenAlex und Freesound verlangen eine Registrierung. Der Schlüssel wird lokal hinterlegt und ausschließlich an die jeweilige Quelle gesendet. OpenAlex ist seit Februar 2026 schlüsselpflichtig; zuvor war der Zugang frei.

Stufe C — nur über optionalen Vermittler (3 Archive). Met Museum, Library of Congress und GDELT senden keine Cross-Origin-Freigabe beziehungsweise drosseln anonyme Zugriffe. Ein beigelegter Vermittler (Cloudflare Worker, rund 100 Zeilen) schließt diese Lücke. Er arbeitet mit einer festen Zielliste, protokolliert nicht und speichert nicht. Ohne ihn läuft die Suite unverändert; die drei Archive erscheinen dann nicht.

Nicht einbindbar bleiben ORF-TVthek und arte. Dort liegt kein technisches, sondern ein rechtliches Hindernis vor — Geosperrung und fehlende Nutzungsrechte. Ein Vermittler ändert daran nichts, weshalb beide im Register als Verweis geführt werden.

4. Lizenzführung

Jeder Treffer erhält eine von drei Stufen, abgeleitet aus den Metadaten des liefernden Archivs.

Abb. 2 — Lizenz-Ampel und Nachweiskette / License traffic light and evidence chain

Zuordnung aus Archiv-Metadaten · Prüfung am Original bleibt erforderlich



Nachweis je Treffer (Export: Markdown · HTML · JSON)

Titel	Urheber	Datum	Archiv	Lizenz	Link	Abrufdatum	Suchbegriff
Die Mappe sammelt Nachweise, keine Inhalte — Metadaten und Verweis, kein kopiertes Material. Export gruppiert nach Ampelstufe und weist Bilanz aus (frei / Namensnennung / nur ansehen).							

Die Ampel ist Orientierung aus Metadaten, keine Rechtsauskunft. Prüfung am Original vor jeder Verwendung.

Stufe	Zuordnung	Nutzung
Grün	CC0, Public Domain	frei, auch kommerziell, ohne Namensnennung
Gelb	CC BY, CC BY-SA	mit Namensnennung; BY-SA fordert gleiche Weitergabe
Rot	Rundfunkrechte, Presse, ungeklärt	belegbar, nicht wiederverwendbar

Die Zuordnung erfolgt maschinell aus Feldern wie `edm:rights` (Europeana), `share_license_status` (Cleveland), `is_public_domain` (Art Institute) oder `LicenseShortName` (Wikimedia). Wo ein Feld fehlt oder unklar bleibt, wird nicht geraten, sondern als ungeklärt geführt.

Ein Punkt verdient Hervorhebung, weil er regelmäßig unterschätzt wird: Freier Zugang und freie Nachnutzung fallen auseinander. Song-MIDIs, kommerzielle Acapellas, Songtexte und Rundfunkarchive sind vielfach abrufbar, ohne nutzbar zu sein. Bei MIDI-Dateien können zwei Rechte nebeneinander bestehen — an der Komposition und am Arrangement. Das Werkzeug führt solche Funde als rot und damit als Beleg, nicht als Material.

5. Funktionsumfang

5.1 Recherche-Station

Die Suche fragt alle aktiven Archive gleichzeitig ab und zeigt je Archiv die Trefferzahl, sodass Ausfälle erkennbar bleiben. Filter bestehen für Medientyp und Lizenzstufe.

Dossiers speichern eine Fragestellung zusammen mit der Auswahl der Archive und lassen sich als Datei weitergeben. Damit wird eine Recherche wiederholbar.

Das Archiv-Register führt 57 Archive auf, auch solche, die das Werkzeug nicht abfragt. Der Zweck liegt in der Auffindbarkeit: Ein Archiv, dessen Existenz unbekannt ist, wird nicht durchsucht. Jeder Eintrag trägt seine Stufe (A, B, C), Region, Medientyp und Lizenzprofil.

Perspektivenvergleich erfolgt über Verweise zu GDELT, Ground News, AllSides und Wikipedia; bei eingerichtetem Vermittler erscheint GDELT unmittelbar in der Trefferliste. Damit lässt sich die Berichtslage zu einem Thema über Sprach- und Ländergrenzen hinweg gegenüberstellen.

Verifikation je Treffer über Rückwärts-Bildsuche (Google Lens, Yandex, Bing, TinEye) sowie Verweis auf das Prüfwerkzeug InVID/WeVerify.

Sammelmappe. Treffer werden mit Titel, Urheber, Datum, Archiv, Lizenz, Verweis, Abrufdatum und verwendetem Suchbegriff gesammelt. Der Export erfolgt als Markdown, HTML oder JSON, nach Lizenzstufe gruppiert und mit Bilanz. Gesammelt werden Nachweise, keine Inhalte — Metadaten und Verweis, kein kopiertes Material.

5.2 Schnitt und Bearbeitung

Wellenformanzeige mit Bereichsauswahl, Verstärkung, Ein- und Ausblendung, Dreiband-Entzerrung, Trenn- und Verzögerungsfilter sowie Tonhöhenänderung mit Tempoangleichung. Der Export erfolgt als WAV.

Tonhöhe und Tempo sind gekoppelt, entsprechend einem Bandmaschinen- oder Plattenspielerverhalten. Eine entkoppelte Dehnung ist nicht enthalten; dafür eignen sich Audacity oder Ardour.

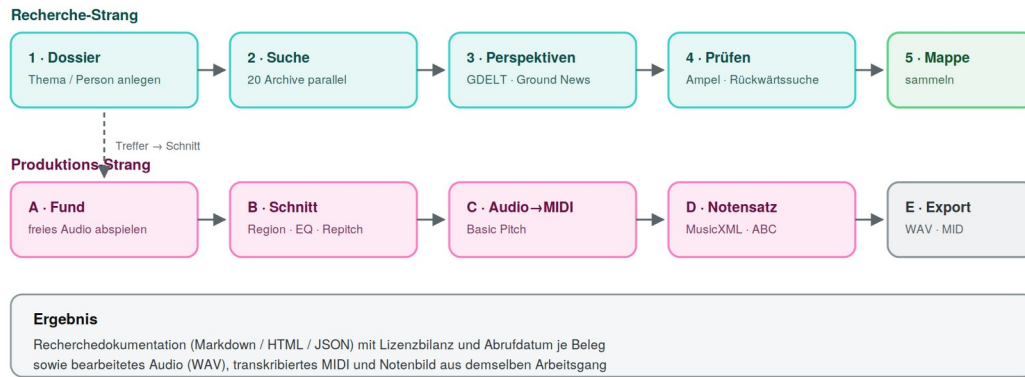
5.3 Noten und MIDI

Audio zu MIDI über Basic Pitch (TensorFlow.js, Apache 2.0), polyphon, mit einstellbarer Ansprechschwelle. **MIDI-Erzeugung** über Magenta.js (MusicVAE, MelodyRNN) für Melodie- und Schlagzeugmuster. **Notensatz** über OpenSheetMusicDisplay für MusicXML und abcjs für ABC-Notation.

Die Umwandlung von MIDI zu Notenbild bleibt eine Näherung. MIDI kennt weder Taktart noch enharmonische Verwechslung; die Zuordnung zu Notenwerten ist eine Auslegung. Das Ergebnis dient als Skizze. Für Notensatz in Veröffentlichungsqualität ist MuseScore vorzuziehen.

Abb. 3 — Arbeitskette / Workflow chain

Von der Fragestellung zur belegten Dokumentation · von Rohmaterial zur Notation



Übergabe zwischen den Strängen erfolgt lokal (IndexedDB); Inhalte verlassen das Gerät nicht.

6. Betrieb

Die vier HTML-Dateien und optional die Vermittler-Datei gehören in ein gemeinsames Verzeichnis; `index.html` bildet den Einstieg. Die Verweise sind relativ.

	Auslieferung über HTTPS	lokaler Aufruf (file://)
Suche in Archiven	vollständig	eingeschränkt
Abspielen, Schnitt, Export	vollständig	vollständig
Mappe, Dossiers, Register	vollständig	vollständig
Audio→MIDI, Notensatz	vollständig	nicht verfügbar

Für den vollen Umfang ist Auslieferung über einen Webserver erforderlich. Browser behandeln lokal geöffnete Dateien restriktiver, und die Bibliotheken für Transkription und Notensatz werden über ein Auslieferungsnetz geladen.

7. Datenschutz

Inhalte werden nicht übertragen. Audiodaten werden lokal entschlüsselt, verarbeitet und exportiert. Analysewerkzeuge oder Zählpixel sind nicht eingebaut.

Nach außen gehen ausschließlich: Suchanfragen an die jeweiligen Archive, Schriftarten von Google Fonts und Programmbibliotheken von jsDelivr. Bei Einsatz des Vermittlers laufen die Anfragen an die drei Archive der Stufe C über diesen; er sieht die Suchbegriffe, speichert sie nicht.

Schlüssel, Dossiers und Sammelmappe verbleiben im Browser.

8. Prüfstand

Vor der Veröffentlichung wurden die Bestandteile geprüft, soweit dies ohne Browser und ohne Zugriff auf die Archive möglich ist.

Prüfung	Umfang	Ergebnis
Auswertungslogik der Schnittstellen	20 Adapter gegen nachgebildete Antworten in echter Feldstruktur	bestanden, 22 Treffer korrekt ausgewertet
Verhalten bei Ausfall	20 Adapter gegen leere und lückenhafte Antworten	ein Absturz gefunden und behoben
Zentrale Absicherung	fehlerhafte Treffer werden vor der Anzeige ausgesondert	bestanden
MIDI-Export	bytewise Rücklesung: Kopf, Spurlänge, variable Zahlenkodierung, Zeitachse	strukturell gültig
WAV-Export	bytewise Rücklesung: Kopf, Abtastwerte, Begrenzung, Kanalverschachtelung	strukturell gültig
Notenumwandlung	Tonleiter, Akkorde, Pausen, Leerfall	bestanden
Dateiwerk	Syntax, Kennungen, Navigation, Formatvariablen, Querverweise	bestanden

Der gefundene Absturz betraf die Radioabfrage bei einer Antwort, die kein Feld erwarteter Form enthielt. Er hätte im Betrieb die laufende Suche unterbrochen. Beides — Fund und Behebung — ist hier vermerkt, weil die Prüfung sonst nicht nachvollziehbar wäre.

Nicht geprüft werden konnte das Verhalten gegen die realen Archive. Die Prüfumgebung hatte keinen Zugang zu diesen Diensten. Ob ein Archiv am Tag des Einsatzes antwortet und seine Zugriffsregeln unverändert sind, zeigt sich erst im Betrieb.

9. Grenzen

1. **Verfügbarkeit der Quellen.** Bei Chronicling America, LibriVox und Openverse war der Browserzugriff in der Erhebung als zu verifizieren vermerkt. Fällt eine Quelle aus, liegt das an deren Server; sie lässt sich einzeln abschalten.
2. **Rundfunkarchive.** MediathekViewWeb liefert Metadaten und Abrufadressen. Die Inhalte bleiben rundfunkrechtlich gebunden, können regional gesperrt sein und werden nach Fristen entfernt.
3. **Zitate.** Wikiquote-Treffer erscheinen als Titel, Kurzkontext und Verweis. Der Zitattext selbst wird nicht übernommen, da Zitate unabhängig von der Lizenz der Sammlung geschützt sein können.
4. **Keine Synchronisierung.** Ohne Serverkomponente bleiben Arbeitsdaten an den jeweiligen Browser gebunden.

5. **Notensatz.** Siehe Abschnitt 5.3 — Skizze, keine Veröffentlichungsqualität.
6. **Tonhöhe und Tempo** sind gekoppelt.

10. Nachnutzung und Erweiterung

Code und Dokumentation stehen unter CC BY 4.0. Eine Quelle wird durch Ergänzung eines Objekts im Feld SOURCES eingebunden:

```
{id:"kurz", label:"Anzeigenname", rg:"Region", media:["audio"],
byok:false, on:true,
run: async q => {
  const d = await jget(`https://api.beispiel.org/search?q=$
{enc(q)}`);
  return (d.results||[]).map(x => ({
    title:x.name, creator:x.author||"", date:x.year||"",
    type:"audio",          // audio | video | image | text | science
    license:"green",       // green | yellow | red
    licName:"CC0", source:"Beispiel", region:"global",
    page:x.url, thumb:x.thumb||"", file:x.audio||null
  }));
}},
```

Voraussetzung ist eine Schnittstelle, die Cross-Origin-Zugriffe erlaubt und JSON liefert. Prüfung vorab:

```
curl -I -H "Origin: https://example.com"
"https://api.beispiel.org/search?q=test"
```

Im Antwortkopf muss Access-Control-Allow-Origin gesetzt sein.

11. Verwendete freie Bibliotheken

Bibliothek	Zweck	Lizenz
Basic Pitch (Spotify)	Audio zu MIDI	Apache 2.0
Magenta.js (Google)	MIDI-Erzeugung	Apache 2.0
OpenSheetMusicDisplay	MusicXML-Darstellung	BSD
abcjs	ABC-Notation	MIT
TensorFlow.js	Ausführung der Modelle	Apache 2.0

Web Audio API, Fetch und IndexedDB sind Bestandteil des Browsers und benötigen keine Bibliothek.

12. Interessenoffenlegung

Der Autor forscht unabhängig, ohne institutionelle Anbindung, privat finanziert und ohne Förder- oder Drittmittel. Daneben betreibt er gewerbliche Tätigkeiten in den Bereichen Fertigung, Audio- und Kreativsoftware sowie Veranstaltungssicherheit.

Für diese Arbeit besteht eine Berührung, die offengelegt sei: Dasselbe Werkzeug wird zusätzlich unter einer eigenen kommerziellen Marke als kostenloses Werkzeug veröffentlicht. Beide

Fassungen beruhen auf demselben Quellcode. Einnahmen entstehen daraus nicht; ein mittelbarer Nutzen durch erhöhte Sichtbarkeit der Marke ist möglich und wird nicht bestritten. Die im Werkzeug behandelten Archive und Bibliotheken stehen in keiner geschäftlichen Beziehung zum Autor.

Da keine institutionelle Anbindung besteht, gibt es auch keine institutionelle Qualitätssicherung. Die Angaben zur Methodik sind darauf angelegt, eine eigenständige Prüfung zu ermöglichen.

Vollständige Erklärung in INTERESSEN OFFENLEGUNG .md, Einsatz von Hilfsmitteln in AI_DISCLOSURE .md.

13. Zitation

Schüller, T. (2026). *Medienrecherche v1.0 (Werkzeuge · Syntagma Forschung): Serverloses Browser-Werkzeug für Recherche, Bearbeitung und Nachweis freier Medien*. Zenodo. [https://doi.org/\[DOI nach Upload eintragen\]](https://doi.org/[DOI nach Upload eintragen])

Herausgegeben im Werkzeuge · Syntagma Forschung. Rückmeldungen zu ausgefallenen Quellen oder fehlenden Archiven sind willkommen, da die Erreichbarkeit von Schnittstellen sich laufend ändert.