

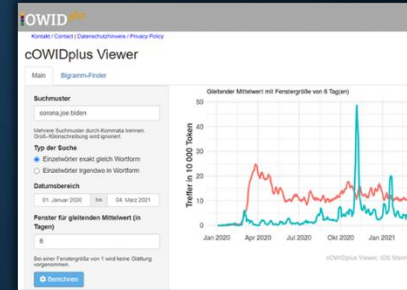
The background of the slide is a photograph of a brick wall and a window frame. The bricks are reddish-brown and arranged in a traditional pattern. A white window frame is visible on the right side. In the background, out of focus, there is a purple metal structure, possibly a playground or exercise equipment, and some greenery.

OWIDplusLIVE : OpenSource Tool zur Exploration ausgewählter Sprachressourcen. Am Beispiel deutschsprachiger Online-Presse zu ‚Impfung‘

**Sascha Wolfer, Jan Oliver Rüdiger, Alexander Koplenig,
Frank Michaelis, Carolin Müller-Spitzer**

Heute

- Einführung in die Datenbasis
- Vorgänger: cOWIDplus Viewer
<https://www.owid.de/plus/cowidplusviewer2020/>
- OWIDplusLIVE
<https://www.owid.de/plus/live-2021/>
- Konkrete Analysebeispiele zu ‚Impfung‘



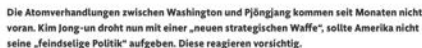
- RSS news feeds
- 13 deutschsprachige Online-Quellen seit Januar 2020
- Stündliche Sammlung (früher: 14-tägig - jetzt: **tägliche Updates**)
- Momentan: **30,1 Millionen Tokens** (Stand 4. März 2021)
- Title and description



Title
Nordkorea: Kim fühlt sich nicht mehr an Atomteststopp gebunden

Nordkorea: Kim fühlt sich nicht mehr an Atomteststopp gebunden

AKTUALISIERT AM 01.01.2020 - 06:40



MERKEN ☆ 4 📧 🔖

Es sei Teil der Bemühungen um gegenseitiges Vertrauen gewesen, die Atomversuche und Tests mit Interkontinentalraketen auszusetzen, wurde Kim am Mittwoch von den staatlichen Medien zitiert. Den Vereinigten Staaten warf der Machthaber vor, sein Land militärisch zu bedrohen und es durch Sanktionen „abzressen“ zu wollen. „Unter diesen Umständen gibt es keinen Grund für uns, länger einseitig an die Verpflichtungen gebunden zu sein.“

Pompeo nennt Entwicklung „tief enttäuschend“

ihlt sich nicht mehr an
teststopp gebunden

M 01.01.2020 - 04:40



Verhandlungen zwischen Washington und Pjöngjang kommen seit
 nicht voran. Kim Jong-un droht nun mit einer „neuen strategischen
 Linie“. Die USA werden gezwungen sein, ihre „feindselige Politik“ aufzugeben. Diese reagieren

cOVIDplus Viewer

Main [Bigramm-Finder](#)

Suchmuster

Mehrere Suchmuster durch Kommata trennen. Groß-/Kleinschreibung wird ignoriert.

Typ der Suche

- ☒ Einzelwörter exakt gleich Wortform
☐ Einzelwörter irgendwo in Wortform

Datumsbereich

bis

Fenster für gleitenden Mittelwert (in Tagen)

Bei einer Fenstergröße von 1 wird keine Glättung vorgenommen.

 Berechnen

 Daten

 PNG

 PDF

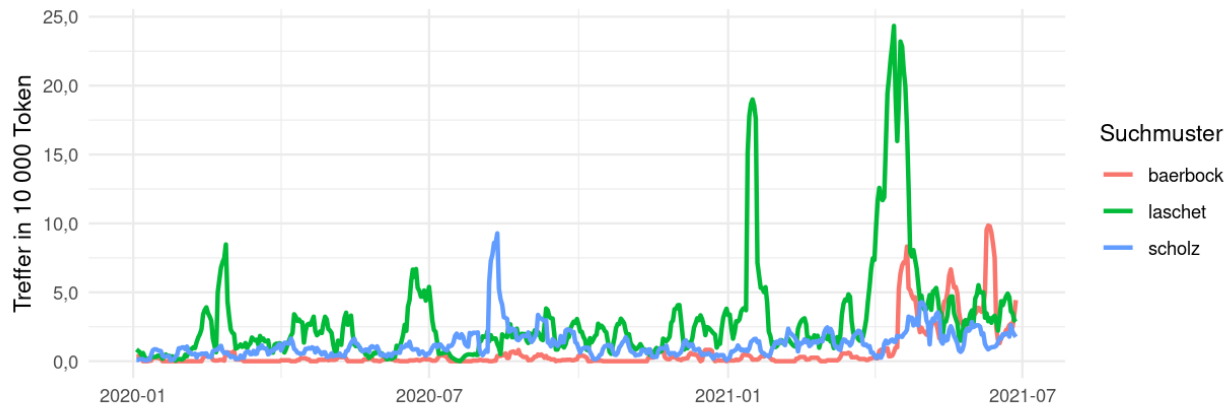
Achtung! Ab dem 5. November 2020 werden Einzelwortformen, die nur einmal am Tag vorkommen, aus Speichergründen nicht mehr in die Berechnung mit einbezogen. Außerdem sind nur Einträge in der Datenbasis, die an mindestens 5 Tagen im Datensatz vorkommen.

Stand der Daten: 30. Juni 2021

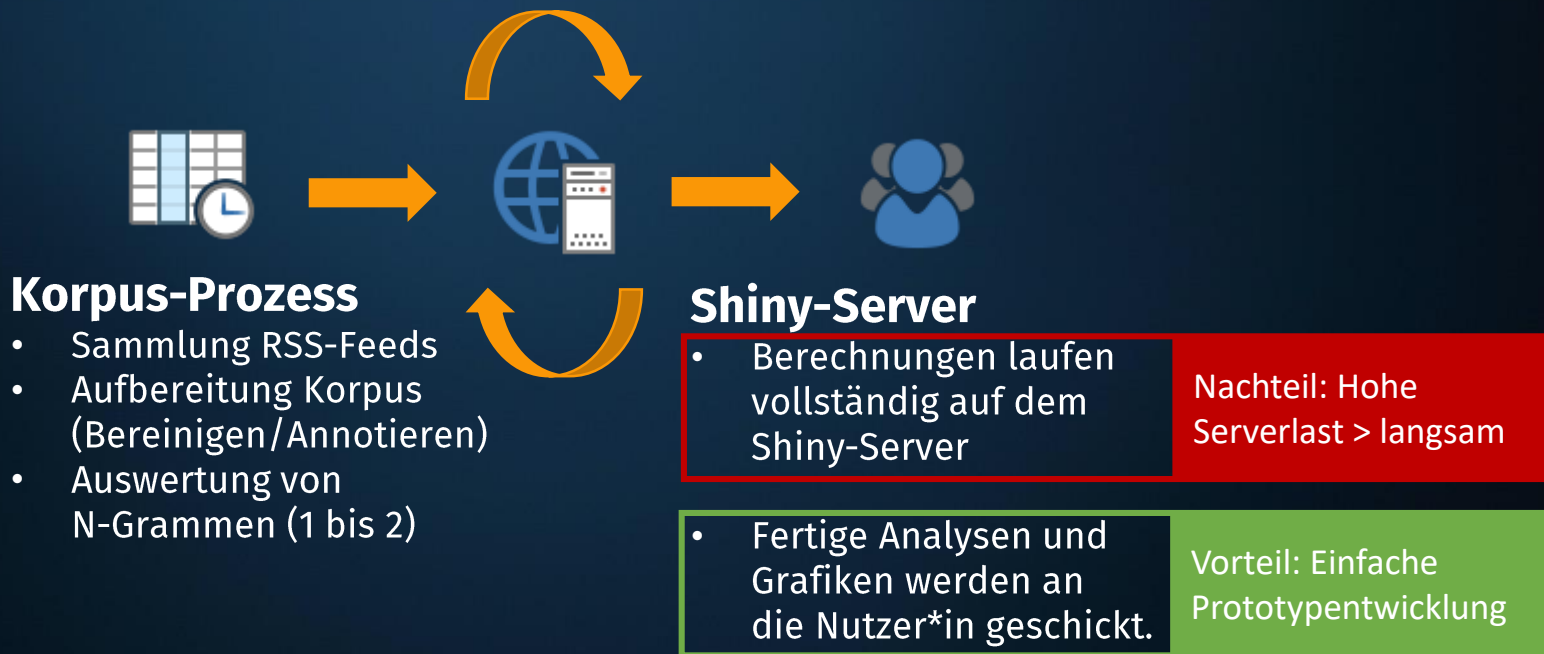
Programmbereich Lexik empirisch & digital, IDS Mannheim

[Englische Version](#)

Gleitender Mittelwert mit Fenstergröße von 6 Tag(en)



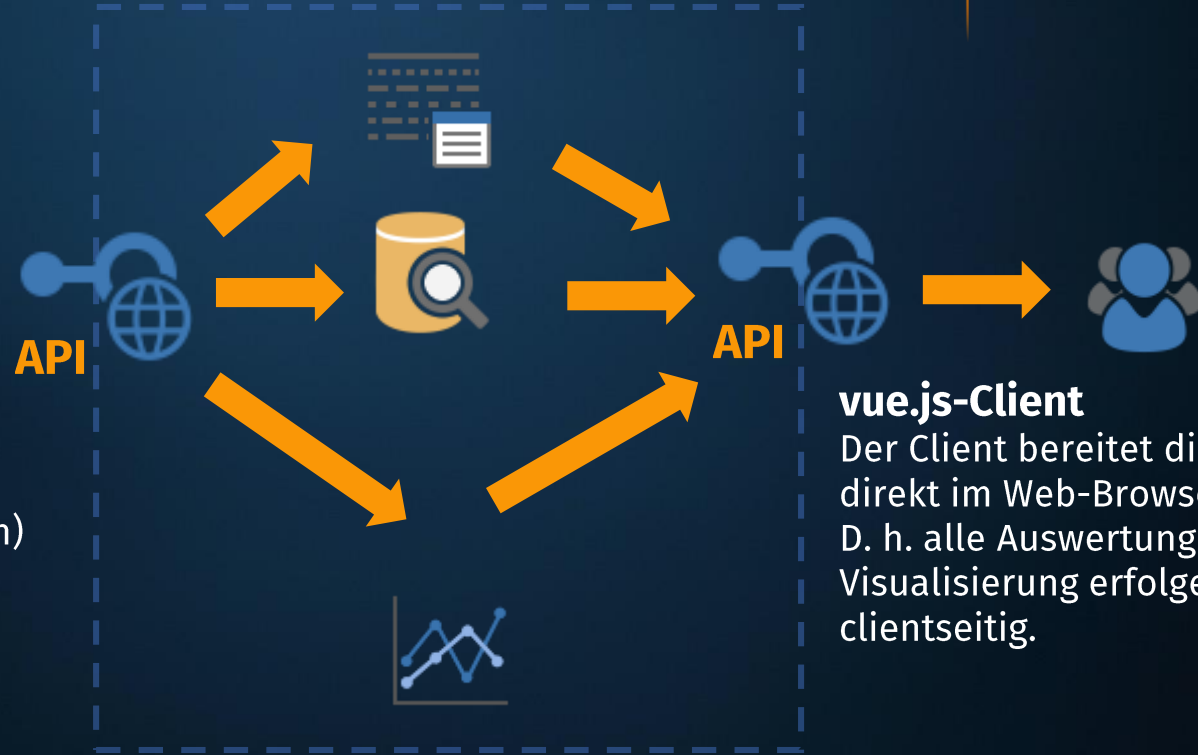
cOVIDplus Viewer, IDS Mannheim





Korpus-Prozess

- Sammlung RSS-Feeds
- Aufbereitung Korpus (Bereinigen/Annotieren)
- Auswertung von N-Grammen (1 bis **3**)



vue.js-Client

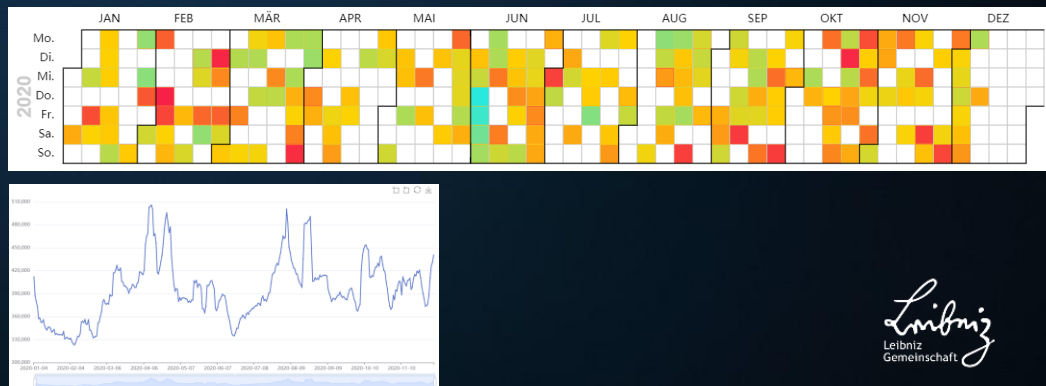
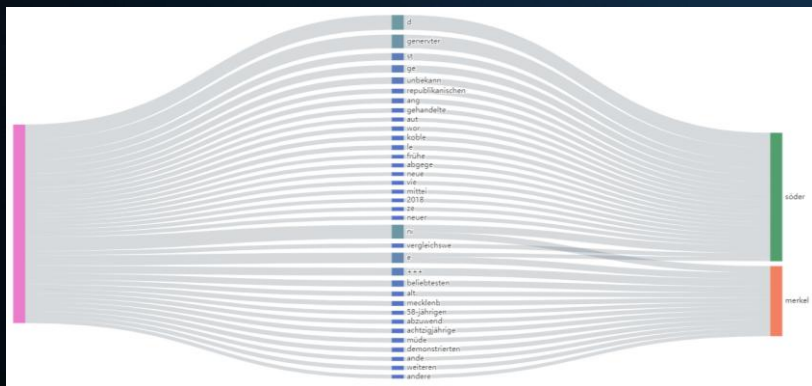
Der Client bereitet die Daten direkt im Web-Browser auf. D. h. alle Auswertungen und Visualisierung erfolgen clientseitig.

OWIDplusLIVE – Key-Features

IDS

LEIBNIZ-INSTITUT FÜR
DEUTSCHE SPRACHE

- Tagesaktuelle Suche nach N(1-3)-Grammen [Frequenzen, Bi- und Tri-Grammen]
- (Wildcard-)Abfragen über die Layer: Wortform, Lemma, Part-of-Speech
- Tiefensuche (Abfrage über mehrere Layer gleichzeitig)
- Drei Visualisierungen (Chart, Kalender, Sankey)
- Visualisierungsoptionen: (absolute/relative) Frequenzen, Granulierung, Glättung
- Zeitreihen können einzeln betrachtet sowie an- und abgewählt werden.
- Export der Visualisierung als Grafik sowie ein Datenexport als JSON/TSV möglich



🔍 SUCHEN: Einfache Suche ⓘ

Suchfenstergröße:

N=1 N=2 N=3

Suchabfrage:

Wortform ▼ Suchausc

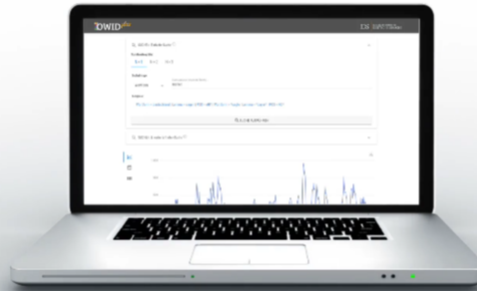
Beispiele (anklicken und ausführen):

Wortform = deutschland | Lemma =

🗑️ ALLE EINGABEN

🔍 SUCHEN: Erweiterte Tiefen-Suche ⓘ

Herzlich willkommen bei OWIDplusLIVE



▶ ⏮ ⏪ 🔊 0:03 / 8:11 ⌵ ⚙️ 🖨️ 🗨️

Dieses Video gibt Ihnen eine kurze Einführung in die Analyse von Token und N-Grammen.

➡️ TUTORIAL BEENDEN

ⓘ



🔍 SUCHEN: Einfache Suche ⓘ



Suchfenstergröße:

N=1

N=2

N=3

Suchabfrage:

Wortform



Suchausdruck (einzelnes Token)

Impfung

Beispiele (anklicken und ausführen):

Wortform = deutschland | Lemma = sagen | POS = ART | Wortform = *sagte | Lemma = *bayer* | POS = KO*



ALLE EINGABEN LÖSCHEN



SUCHE AUSFÜHREN

🔍 SUCHEN: Erweiterte Tiefen-Suche ⓘ



Q SUCHEN: Einfache Suche ⓘ

Suchfenstergröße:

1

N=1

N=2

N=3

Suchabfrage:

2

Wortform

3

Suchausdruck (einzelnes Token)

Impfung

Beispiele (anklicken und ausführen):

?

Wortform = deutschland | Lemma = sagen | POS = ART | Wortform = *sagte | Lemma = *bayer* | POS = KO*



ALLE EINGABEN LÖSCHEN

4



SUCHE AUSFÜHREN

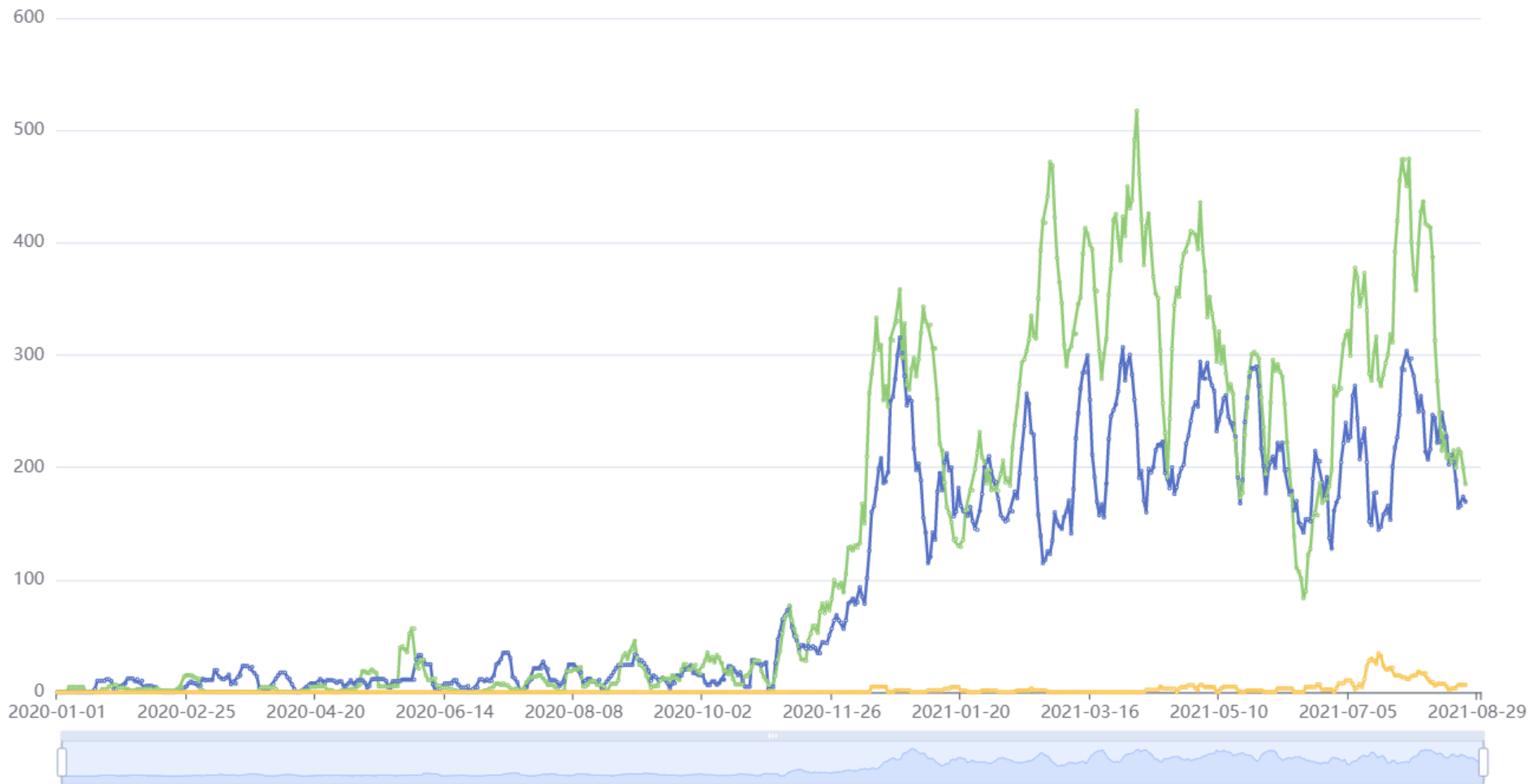
SUCHEN: Erweiterte Tiefen-Suche ⓘ



○ (N = 1): [1. Wortform] = impfung [1 / 1]

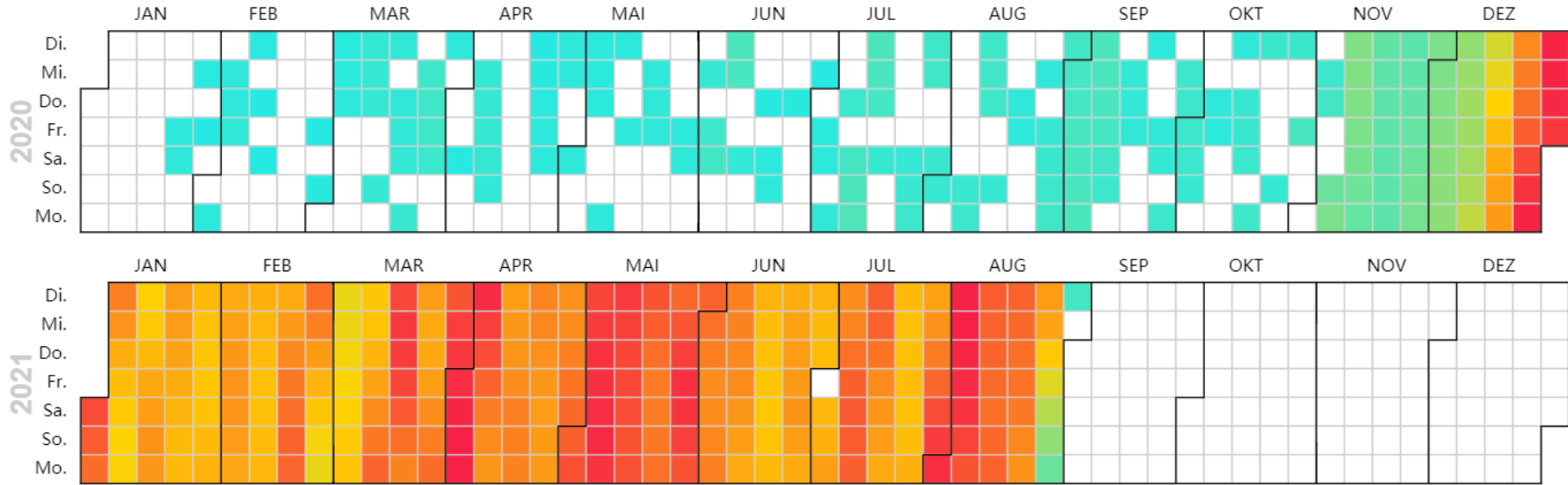


—○— (N = 1): [1. Wortform] = impfung [1 / 1] —○— (N = 1): [1. Wortform] = impfen [3 / 3] —○— (N = 1): [1. Wortform] = impfverweigerer [1 / 1]



○ (N = 1): [1. Wortform] = impfung [1 / 1]



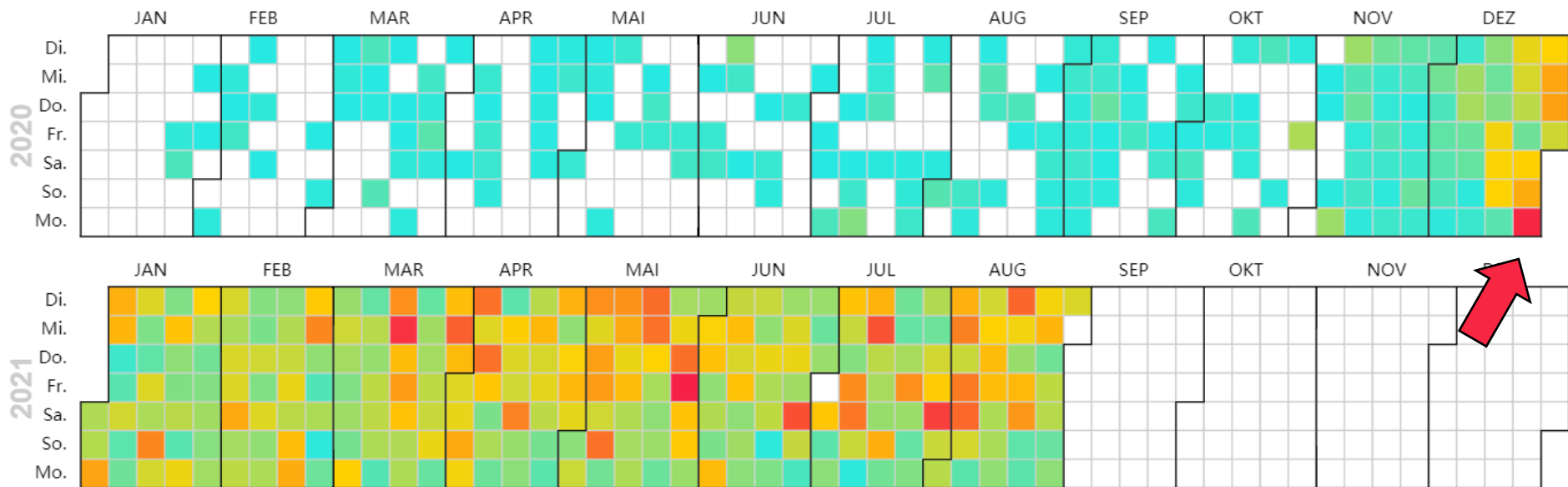


ANPASSEN: Einstellungen für Visualisierung ändern ⓘ

☒ Relative Frequenz

Granulierung: ☐ Tag ☐ Woche ☐ Monat ☐ Quartal ☐ Jahr

Glättung: 1 7 14 21 28 31



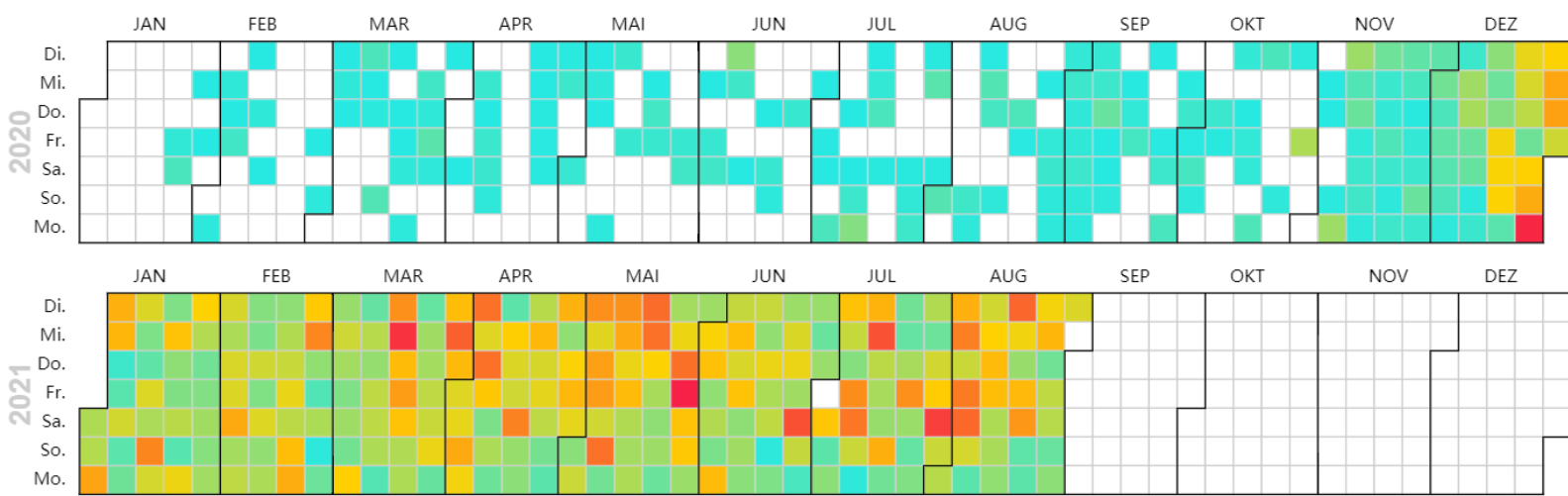
ANPASSEN: Einstellungen für Visualisierung ändern ⓘ

Relative Frequenz

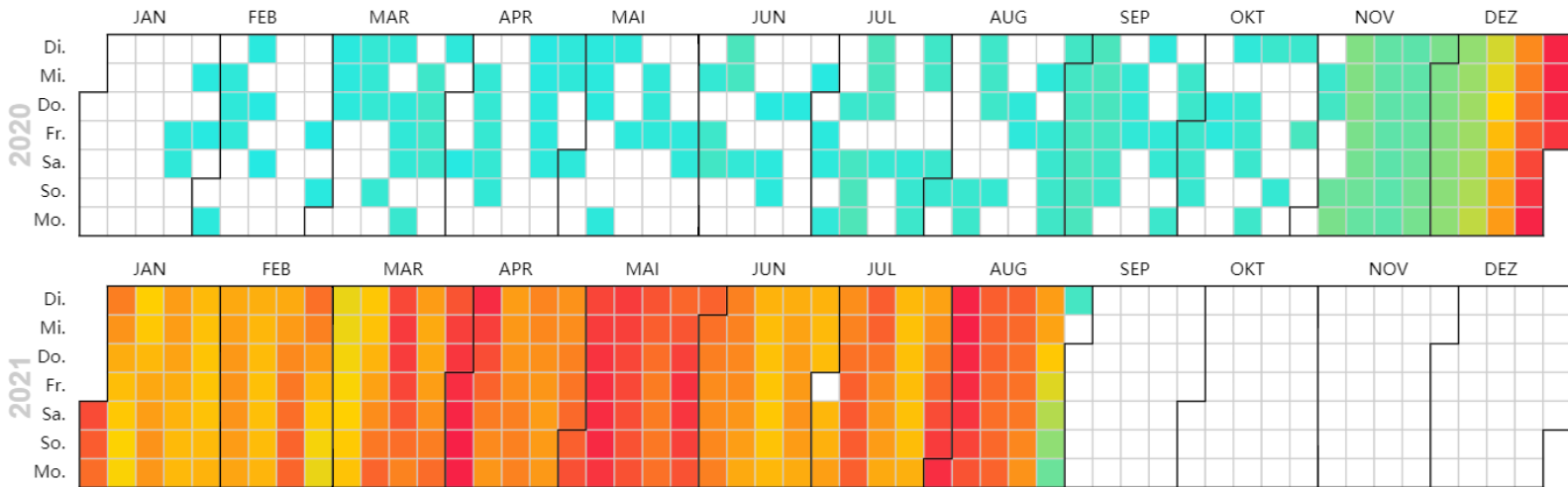
Granulierung: Tag Woche Monat Quartal Jahr

Glättung: 1 7 14 21 28 31

Rohdaten (ohne Glättung)



Gleitender Mittelwert (7 Tage)



🔍 SUCHEN: Einfache Suche ⓘ

Suchfenstergröße:

N=1 N=2 N=3

Suchabfrage:

Wortform Suchausdruck (an 1. Position)... Impfung Wortform Suchausdruck (an 2. Position)... gegen POS ⓘ Suchausdruck (an 3. Position)... NN

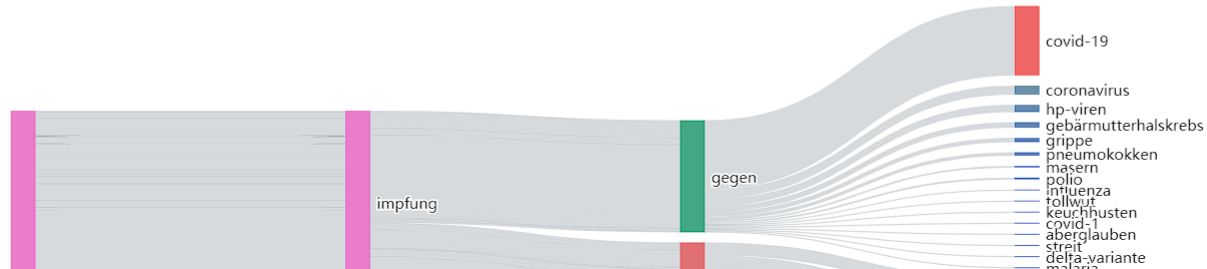
Beispiele (anklicken und ausführen):

[POS = ART / Lemma = unglaublich / POS = NN](#) | [Wortform = merkel / POS = KON / POS = NE](#) | [POS = APPR / Lemma = menschlich / POS = NN](#)

🗑️ ALLE EINGABEN LÖSCHEN

🔍 SUCHE AUSFÜHREN

🔍 SUCHEN: Erweiterte Tiefen-Suche ⓘ



Q SUCHEN: Einfache Suche ⓘ



Suchfenstergröße:

N=1 **1** → N=3

Suchabfrage:

2 → Wortform

Suchausdruck (an 1. Position)...

Impfung

Wortform

Suchausdruck (an 2. Position)...

gegen

3 → POS

Suchausdruck (an 3. Position)...

NN

Beispiele (anklicken und ausführen):

POS = ART / Lemma = unglaublich / POS = NN | Wortform = merkel / POS = KON / POS = NE | POS = APPR / Lemma = menschlich / POS = NN



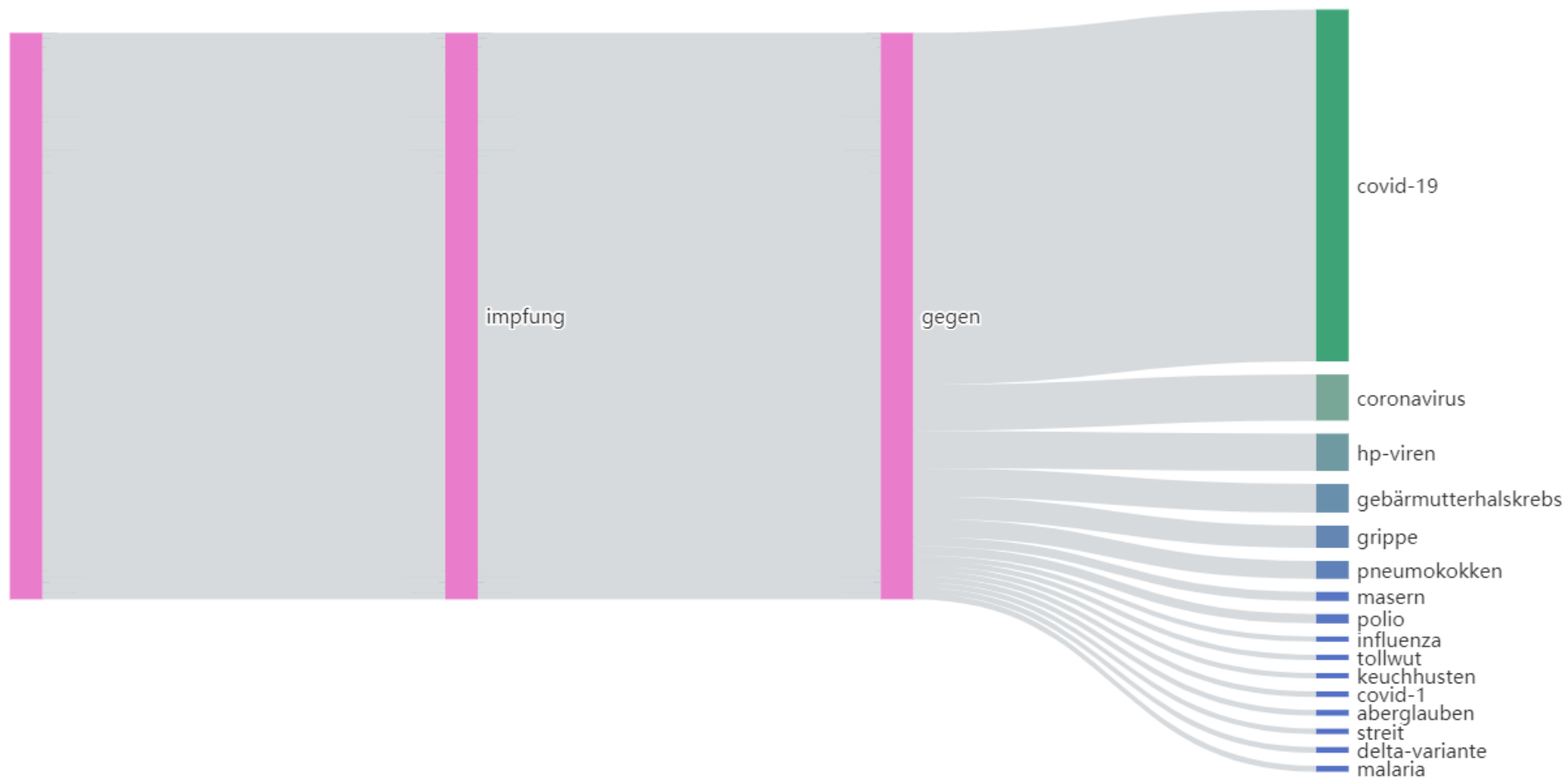
ALLE EINGABEN LÖSCHEN

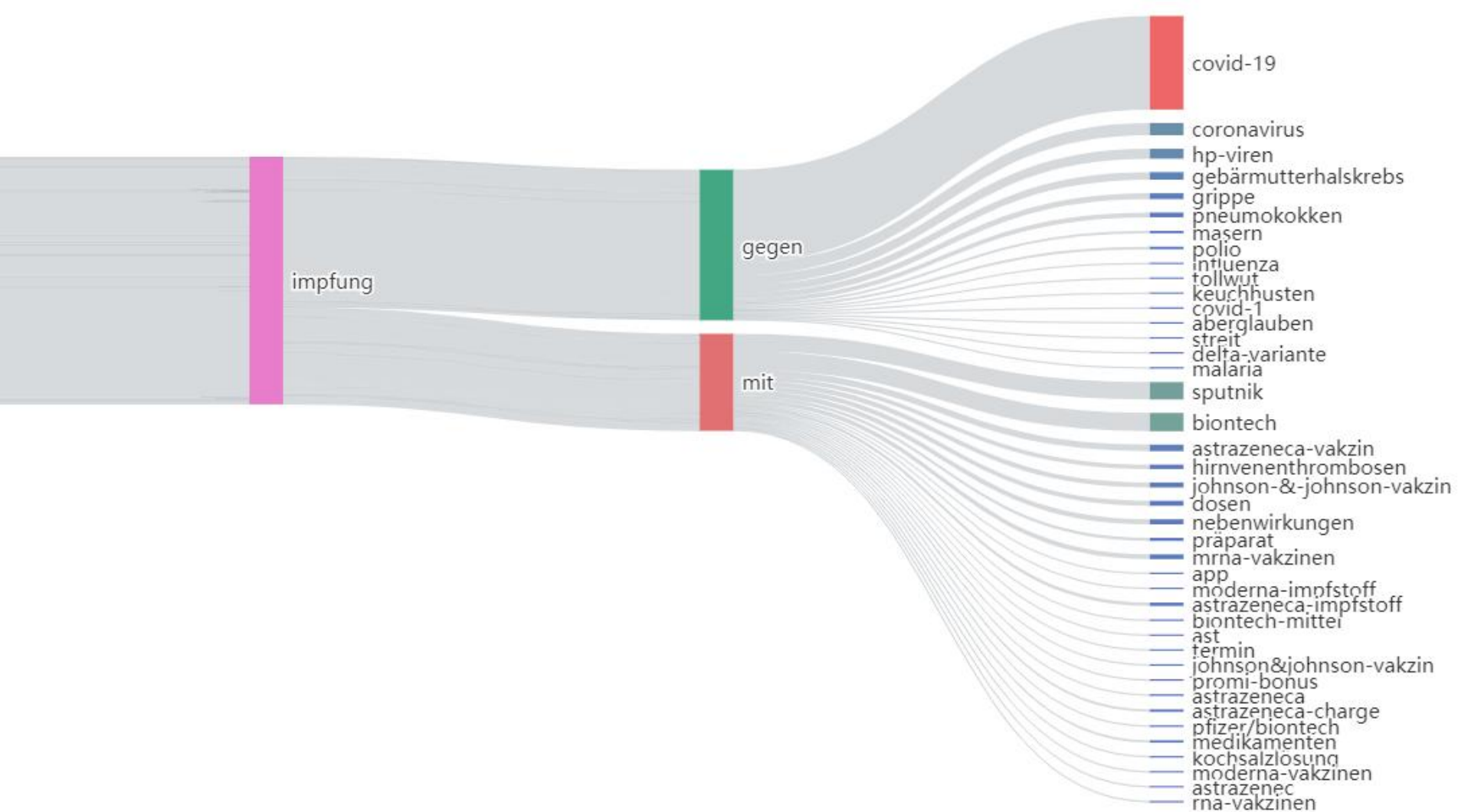


SUCHE AUSFÜHREN

Q SUCHEN: Erweiterte Tiefen-Suche ⓘ







- Tagesaktuelle Suche nach N(1-3)-Grammen [Frequenzen, Bi- und Tri-Grammen]
- (Wildcard-)Abfragen über die Layer: Wortform, Lemma, Part-of-Speech
- Tiefensuche (Abfrage über mehrere Layer gleichzeitig)
- Drei Visualisierungen (Chart, Kalender, Sankey)
- Visualisierungsoptionen: (absolute/relative) Frequenzen, Granulierung, Glättung
- Zeitreihen können einzeln betrachtet sowie an- und abgewählt werden.
- Export der Visualisierung als Grafik sowie ein Datenexport als JSON/TSV möglich

Fragen & DEMO

