



# Teilautomatisierte Erschließung von digitalen Daten mittels Python-Skript

Isabell Schönecker  
Niedersächsisches Landesarchiv  
Abteilung Zentrale Dienste



# Aufbau Cloudplattform ÖGD-Intern

The screenshot displays the user interface of the 'ÖGD-Intern' web application. The top navigation bar is red and contains the logo 'ÖGD-Intern', the text 'Öffentlicher Gesundheitsdienst in Niedersachsen', and the logo for 'Niedersächsisches Landesgesundheitsamt NLGA'. A red 'Abmelden' button is located in the top right corner.

The left sidebar is a vertical menu with the following items: 'Home', 'COVID-19', 'Infektionen & Hygiene', 'Umwelt und Gesundheit' (highlighted in red), 'Erlasse zur TrinkwV', 'Erlasse EU-Badegewässer', 'Erlasse zur Kleinbadeteichen', 'Aabel-Projektbericht', 'GBE', 'Regionales Krebsgeschehen', 'ZGI', and 'Rechtsprechung'. Above the sidebar, it indicates 'Aktuelle Ausgabe des Inf.fo: 25/2023'.

The main content area is titled 'Umwelt und Gesundheit' and contains the following text and links:

- Hier finden Sie Informationen zu Themen aus dem Bereich "Umwelt & Gesundheit".
- [Erlasse zur TrinkwV](#)  
Erlasse und Dokumente des MS zur Trinkwasserverordnung 2001.
- [Erlasse EU-Badegewässer](#)  
Erlasse und Dokumente des MS zu EU-Badegewässern.
- [Erlasse zur Kleinbadeteichen](#)  
Erlasse und Empfehlungen des MS zur Überwachung von Kleinbadeteichen mit biologischer und mechanischer Aufbereitung.
- [Aabel-Projektbericht](#)  
Aabel-Endbericht zum Projekt Aabel "Atemwegserkrankungen und Allergien bei Einschulungskindern in einer ländlichen Region"



## Bewertung ÖGD-Intern

- inhaltliche Bewertung der Daten durch fachlich zuständige Person
  - Daten redundant in verschiedenen Formen verfügbar
    - Frontend = Website
    - Backend = Fileablage
- Entscheidung treffen



Ausgabe	Themen	Anlagen
<a href="#">inf.fo 22/2006</a>	Informationen zu Masern-Ausbruechen	<a href="#">NLGA-Merkblatt Masern</a> <a href="#">Arztinfo Masern, aktualisierte Fassung vom 1.6.</a> <a href="#">Masern-Erhebungsbogen</a>
<a href="#">inf.fo 21/2006</a>	RKI-Infobrief zur Erfassung von Listeriose-Faellen	<a href="#">RKI-Infobrief 14: Erfassung von Listeriose-Fällen</a>
<a href="#">inf.fo 20/2006</a>	Testwoche zur intensivierten Infektions-Surveillance waehrend der Fussball-Weltmeisterschaft 2006 Zur Uebermittlung der IfSG-Melddaten	

## Übernahme

- Aufbau SIP = Struktur der Plattform
- aber: teilweise wurden zusammengehörige Sachverhalte in getrennten Ordnern übergeben („Inf.fo Archiv“)

Home

COVID-19

Infektionen & Hygiene

Meldewesen

Lebensmittelbedingte  
Krankheitsausbrüche

Infektionsschutz/Flucht und  
Migration

Prostituiertenschutzgesetz

Das niedersächsische  
Hygienesiegel für Alten-  
und Pflegeeinrichtungen  
(NiSiP)

Fachausschuss  
Infektionsschutz

AG Krankenhaushygiene/  
Begehungen

Hepatitis A-Leitfaden

Inf.fo Archiv



# AIP-Bildung und Erschließung ÖGD-Intern

- Ziel für die AIP-Bildung des übernommenen „Inf.fo Archivs“:
  - pro Jahrgang ein AIP, mit separatem Unterordner für jeden Bericht und die dazugehörigen Anhänge
- Problem:
  - übernommene Ordnerstruktur ermöglicht keine vollständige Zuordnung von Bericht und Anhängen



*schematische Abbildung*



 [inf.fo 22/2006](#)

Informationen zu Masern-Ausbruechen

 [NLGA-Merkblatt Masern](#)

 [Arztinfo Masern, aktualisierte Fassung vom 1.6.](#)

 [Masern-Erhebungsbogen](#)

 [inf.fo 21/2006](#)

RKI-Infobrief zur Erfassung von Listeriose-Faellen

 [RKI-Infobrief 14: Erfassung von Listeriose-Fällen](#)

```
<td nowrap class='p-2'>  
  <a href="file/inffo0621.pdf" inf.fo 21/2006</a>  
</td>
```

```
<td class='p-2'>  
  <ul class='m-0 p-0 list-group'>  
    <li class="list-group-item p-1">  
      RKI-Infobrief zur Erfassung von Listeriose-Faellen  
    </li>  
  </ul>  
</td>
```

```
<td class='p-2'>  
  <ul class='m-0 p-0 list-group'>  
    <li class="list-group-item p-1">  
      <a href="attachment:ibrki14.pdf">RKI-Infobrief 14: Erfassung von Listeriose-Fällen</a>  
    </li>  
  </ul>  
</td>
```



# Transformation der html-Datei

- Output des Python-Skripts in csv

806	inffo0624	Gehaeuftes Auftreten von HUS-Faellen durch EHEC O157	HUS_06-2006_FB.doc H
807	inffo0623	Startschuss fuer die intensivierete Infektions-Surveillance wa	GAE-Vortagesbericht06
808	inffo0622	Informationen zu Masern-Ausbruechen	MasernMB06_05_261.p
809	inffo0621	RKI-Infobrief zur Erfassung von Listeriose-Faellen	ibrki14.pdf

- Auszug aus der Erschließungsexcel, die für den Ingest mit dem DIMAG IngestTool (DIT) genutzt wird

Signatur	Titel (nicht wiederholbar)	Laufzeit (nicht wiederholbar)	Enthält (wiederholbar)
Acc. 2023_901 Nr. 33	Infektionsepidemiologisches Forum (inf.fo) Jahrgang 2006	2006	Norddeutschland Lebensmittel-Warnung: Salmonella in Teewurst#inffo0623: Startschuss für die intensivierete Infektions-Surveillance während der Fussball- Weltmeisterschaft 2006 Bekämpfung von Masern#inffo0622: Informationen zu Masern-Ausbrüchen #inffo0621: RKI-Infobrief zur Erfassung von Listeriose- Fällen#inffo0620: Testwoche zur intensivierten Infektions-Surveillance während der Fussball-Weltmeisterschaft 2006 Zur Übermittlung der IfSG- Meldedaten#inffo0619: Merkblatt zur Erlaubnis für Tätigkeiten mit



# Gliederung im Archivinformationssystem

Archive in Niedersachsen und Bremen

Nds. Landesarchiv, Abt. Hannover

Nds. 345 Niedersächsisches

Landesgesundheitsamt

6 ÖGD-Intern (Webplattform des Öffentlichen Gesundheitsdienstes in Niedersachsen)

6.1 COVID-19 (SARS Cov 2)

6.2 Infektionen und Hygiene

6.3 Umwelt und Gesundheit

6.4 Gesundheitsberichterstattung (GBE)

6.5 Zentrum für Gesundheits- und Infektionsschutz (ZGI)

6.6 Infektionsepidemiologisches Forum (inf.fo)

NLA HA > Nds. 345 > Infektionsepidemiologisches Forum (inf.fo)

1 / 2

	NLA HA, Nds. 345, Acc. 2023/901 Nr. 32	Infektionsepidemiologisches Forum (inf.fo) Jahrgang 2005
	NLA HA, Nds. 345, Acc. 2023/901 Nr. 33	Infektionsepidemiologisches Forum (inf.fo) Jahrgang 2006
	NLA HA, Nds. 345, Acc. 2023/901 Nr. 34	Infektionsepidemiologisches Forum (inf.fo) Jahrgang 2007
	NLA HA, Nds. 345, Acc. 2023/901 Nr. 35	Infektionsepidemiologisches Forum (inf.fo) Jahrgang 2008



# Erschließung im Archivinformationssystem

NLA HA Nds. 345 Acc. 2023/901 Nr. 32

Neue Repräsentation Exportieren Kontext an:

Bearbeiten | Löschen | Verschieben | Drucken | In den Bestellkorb | Merken | Verlinken | Versenden | Verbo

## Beschreibung - Repräsentationen - Eigenschaften

### Beschreibung: Verzeichnung

#### Identifikation

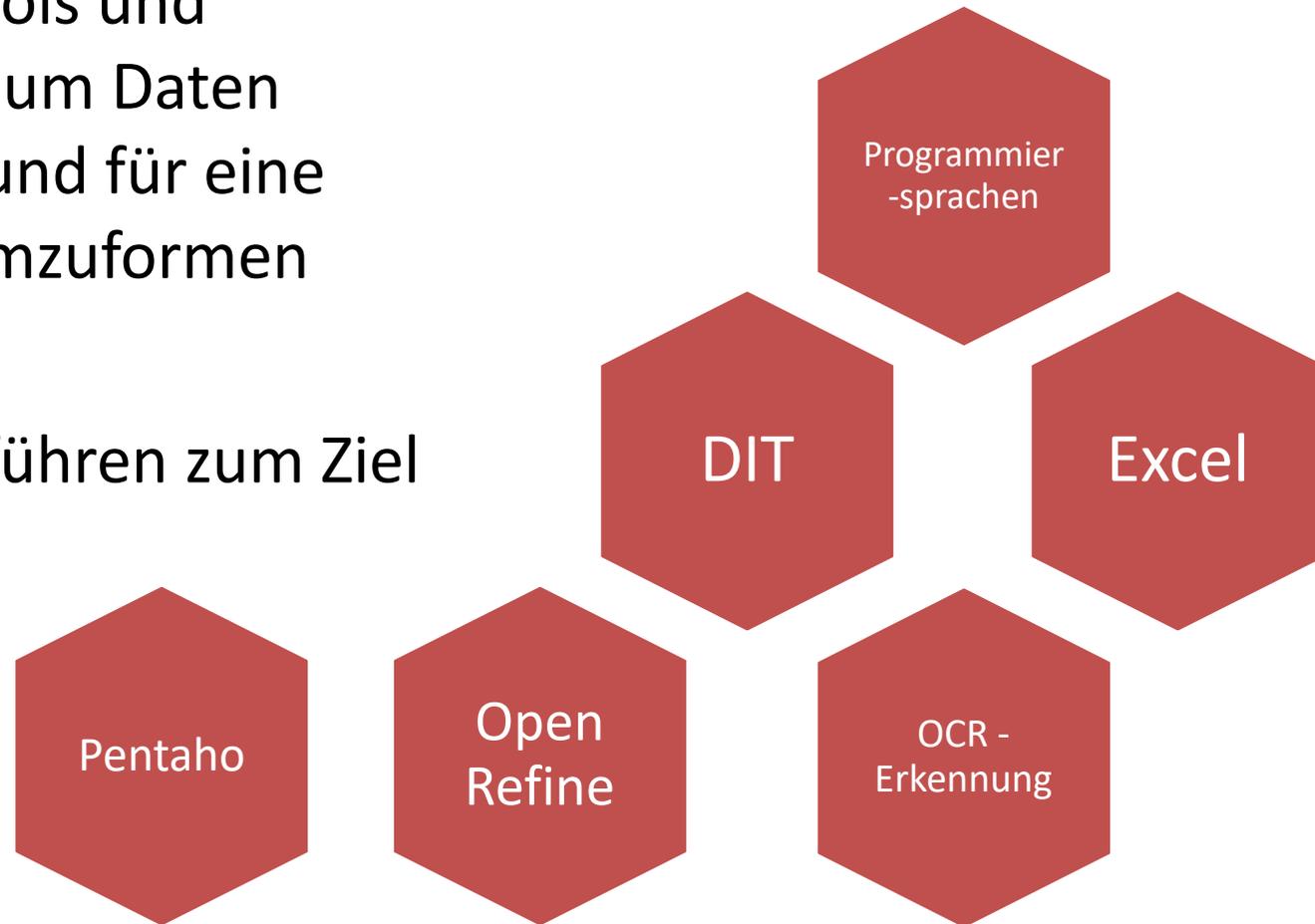
<i>Titel</i>	Infektionsepidemiologisches Forum (inf.fo) Jahrgang 2005
<i>Laufzeit</i>	2005
<i>Enthält</i>	inffo0552: Jahresrückblick 2005 inffo0551: Laborsurveillance von Meningokokken inffo0550: In eigener Sache: NLGA begeht 10-jähriges Jubiläum inffo0549: Mehr Meldefälle durch Campylobacter als durch Salmonellosen in 2005 inffo0548: Tularämie - Hasenpest inffo0547: Norovirusausbrüche durch Himbeeren inffo0546: Probenmaterial für virologische Untersuchungen; S. Goldcoast bei Mallorcareisenden; Norovirusausbruch nach Betriebsversammlung inffo0545: Verdachtsfälle von Wundbotulismus bei intravenös drogenabhängigen Patienten inffo0544: Übermittelte Todesfälle; STEC in Salami inffo0543: Salmonella Virchow aus Hammamed, Tunesien; Salmonella Altona in Sousse, Tunesien; Lebensmittel-Warnung: Noroviren in Himbeeren/Himbeergries aus Polen



# Tiefenerschließung ermöglichen – Wie?

Beispiele für Tools und  
Möglichkeiten, um Daten  
aufzubereiten und für eine  
Erschließung umzuformen

→ viele Wege führen zum Ziel





# Tiefenerschließung ermöglichen – Wie?

- z.B. gesamte Dokumentinhalte mittels OCR erkennen
  - wäre eine generelle OCR Erkennung von schriftbasiertem Archivgut und Bereitstellung für die Recherche möglich/sinnvoll?
- Automatische Transkripterstellung
- Nutzung von KI zur Interpretation audiovisueller Medien

## Frage:

- generierte Informationen Nutzenden über Erschließungs-  
informationen bereitstellen oder „nur“ mit dem Archivgut  
aushändigen?
  - für Strukturdaten oder Normdaten verwenden?



# Tiefenerschließung ermöglichen – Warum?

- auf digitale Daten können schnell Maßnahmen angewendet werden
- unklar, wie Nutzende künftig Archivalien auswerten und interpretieren wollen
  - je mehr Stichworte/ je tiefer die Erschließung, desto besser?
  - Verhältnis von Nutzen und Aufwand
- Differenzierung zu ähnlichem Archivgut des gleichen Themas schwer möglich

NLA HA, Nds. 345,  
Acc. 2023/901 Nr. 33

Infektionsepidemiologisches Forum (inf.fo) Jahrgang  
2006

NLA HA, Nds. 345,  
Acc. 2023/901 Nr. 34

Infektionsepidemiologisches Forum (inf.fo) Jahrgang  
2007



**Kontakt:**  
**Niedersächsisches Landesarchiv**  
**Abteilung Zentrale Dienste**

Isabell Schönecker

[isabell.schoenecker@nla.niedersachsen.de](mailto:isabell.schoenecker@nla.niedersachsen.de)

