

**Retrodigitalisierungsprojekt im
Parlamentsarchiv des Deutschen Bundestages:**

Die Gesetzesdokumentation

Stefan Hauff-Hartig

**Deutscher Bundestag
Referat ID 2, Parlamentsarchiv**





Übersicht, Inhalt

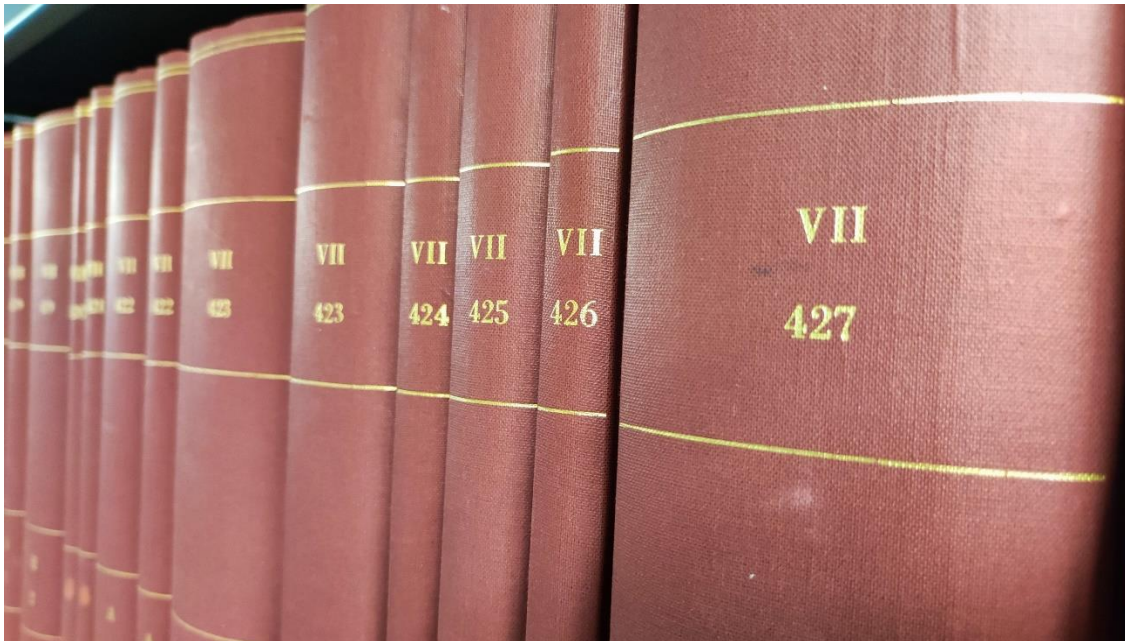
Was ist eine Gesetzesdokumentation	3
Warum digitalisieren?	5
Wer digitalisiert - und wie?.....	7
Was machen wir?	11
Wie sieht das Ergebnis aus?	13
Worin liegen Stärken und Schwächen?.....	21
Was ist noch zu erwähnen?.....	22
Raum für Fragen	23

Was ist eine Gesetzesdokumentation



- Unterlagen des Deutschen Bundestages (BT) und des Bundesrates (BR), die im Gesetzgebungsverfahren entstanden sind
- Überblick über die Entstehung eines Gesetzes und die dazu durchgeführten Beratungen (z.B. Ausschusssitzungen, Plenarsitzungen von BT und BR) – veröffentlichte und nichtveröffentlichte Dokumente
 - für alle verkündeten Bundesgesetze
 - für Gesetzesvorhaben, die nicht verabschiedet wurden

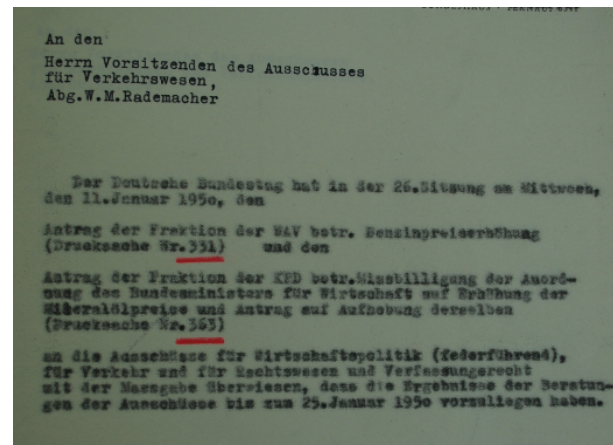
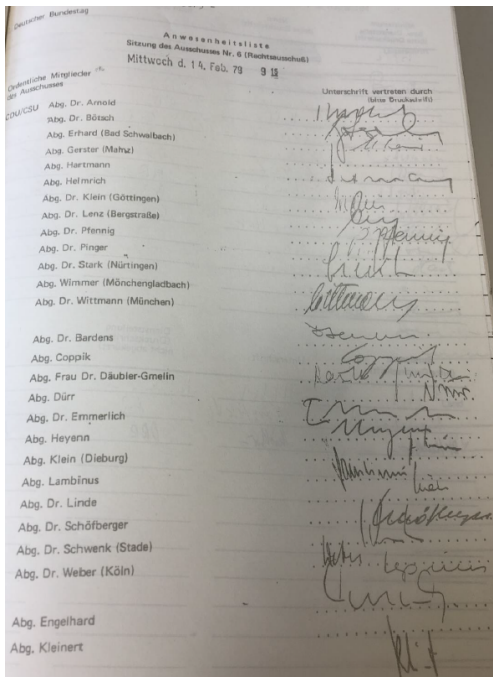
Was ist eine Gesetzesdokumentation? (2)



- seit 1949 (= 1. Wahlperiode inkl. Materialien zum Grundgesetz) in doppelter Ausführung
- Stand 31.12.2021:
 - 1.047 lfd. Meter Gesetzesdokumentationen
 - 10.811 Dokumentationen
 - 25.712 Dokumentationsbände

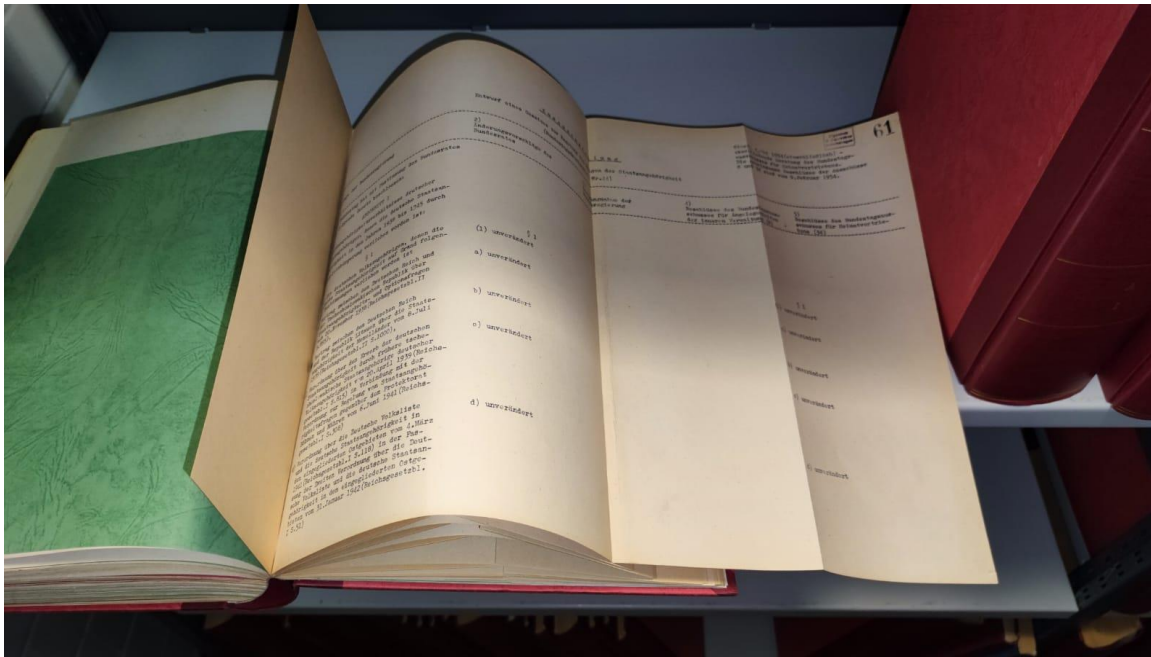


Warum digitalisieren?



Ausgangspunkt: Projekt der Bestandserhaltung, da ältere Bestände teilweise in kritischem Erhaltungszustand

Warum digitalisieren? (2)



- hoher Stellenwert der Materialien
 - für parlamentarische Arbeit
 - für zeitgeschichtliche Forschung
- Volltextsuche möglich
- schnellere Bearbeitung von Anfragen
- einfachere Bereitstellung von Materialien



Wer digitalisiert - und wie?



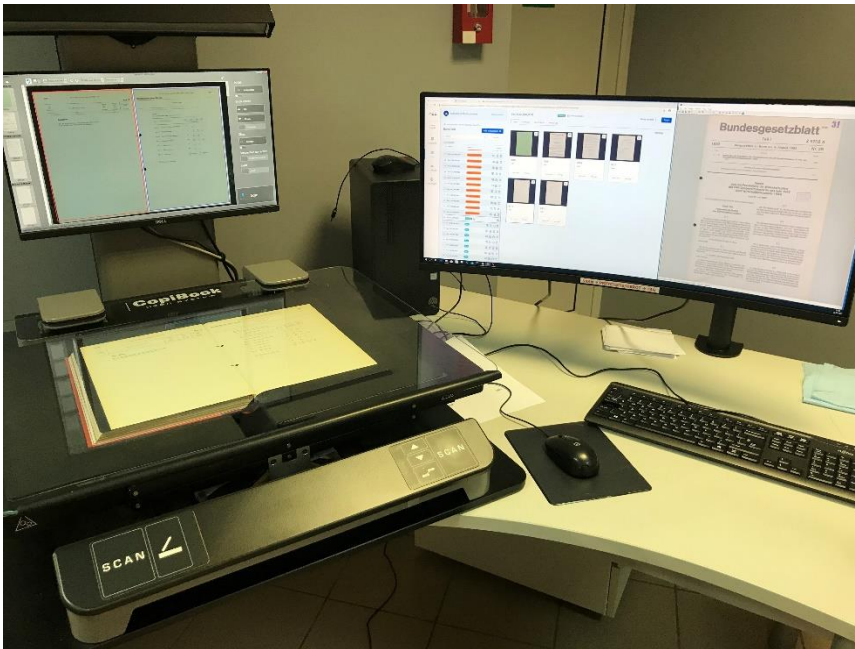
- Projektumfang: Gesetzesmaterialien der 1. bis 16. WP (12.000 Einzelbände mit ca. 5 Mio S.)
- Zeitabdeckung: 1949 bis 2009

Wer digitalisiert - und wie? (2)



- Dienstleister: Bundesdruckerei (BDr)
- Umsetzung durch BDr-Tochterfirma iNCO Sp. z o.o.
- Sitz in Gorzów Wielkopolski (Westpolen), ca. 150 km westlich von Berlin
- umfangreiche Vorbereitungen zur Erstellung der Leistungsbeschreibung und Vertragsvorbereitung notwendig (Vorgaben zur Zielstruktur, technische Parameter, ...)
- Vertragsabschluss: 04/2019, erste Lieferung von Scans: Anfang 2020
- seitdem: alle 3-4 Monate Dateiübergabe und Abholung sowie Rücklieferung von Bänden

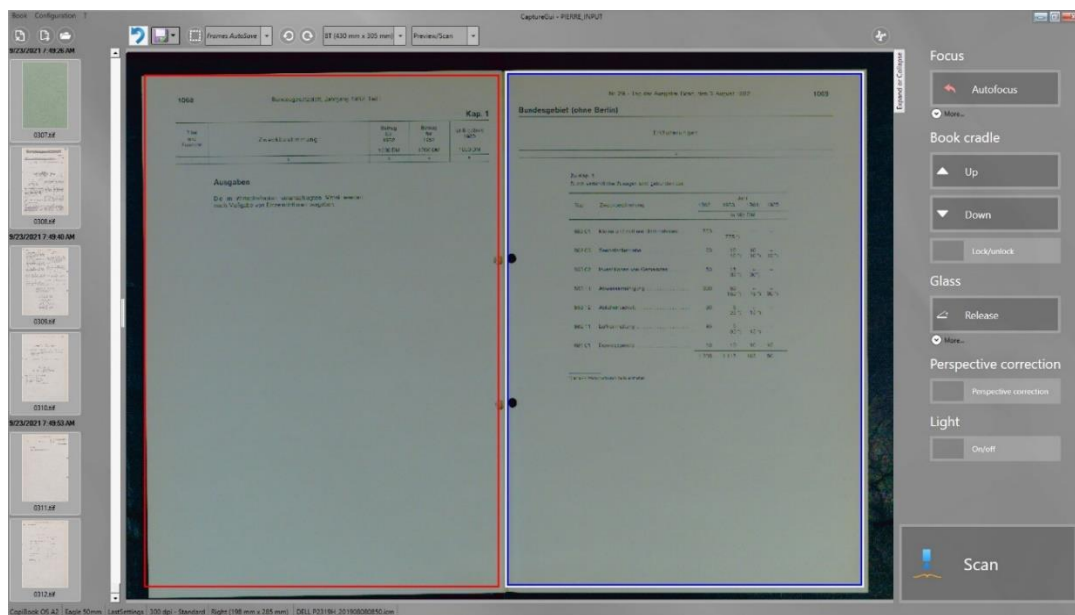
Wer digitalisiert - und wie? (3)



- Scanbereich:
 - 2 Scanstationen (CopiBook i2S) in zwei Schichten, zusätzlich: Einzugs-Scanner für Großformate
 - TIFF-Dateien, aus denen PDFs generiert werden
 - hoher Qualitäts- und Dokumentationsaufwand

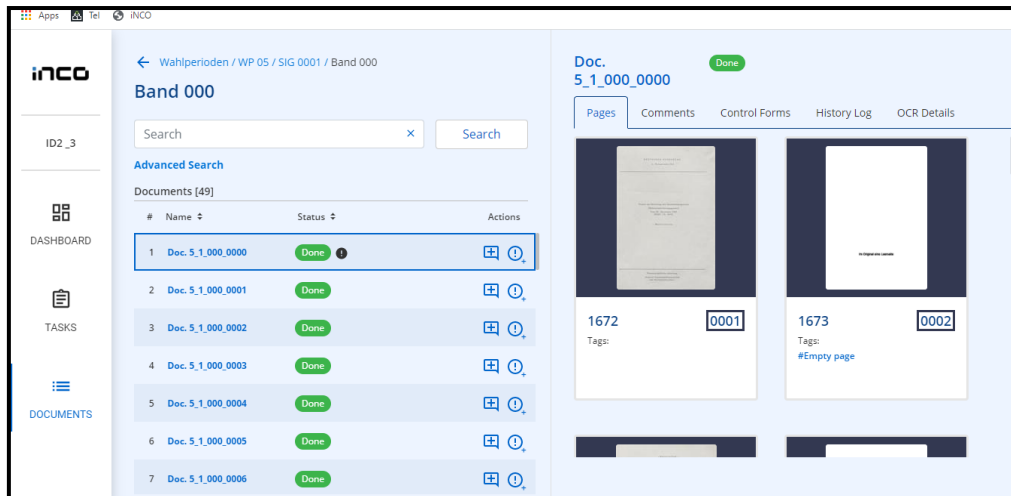


Wer digitalisiert - und wie? (4)



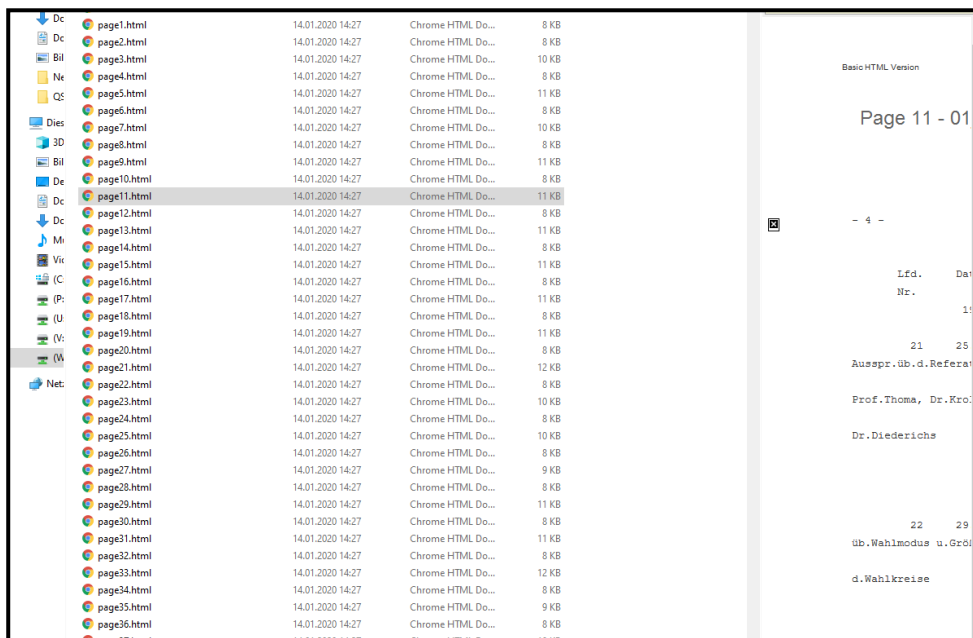
- OCR-Bereich:
 - ABBYY FineReader
 - umfangreiche manuelle Nachbearbeitung, deshalb hohe OCR-Qualität (wichtig v.a. bei Namen von Abgeordneten)

Was machen wir?



- Projektteam des Parlamentsarchivs kann durch Zugriff auf Workflowsystem des Dienstleisters bereits freigegebene Dokumente prüfen
- Qualitätssicherung zu Metadaten, Struktur, Reihenfolge, Vollständigkeit, Ausrichtung, Lesbarkeit, OCR-Präzision → Korrekturen durch DL erfolgen noch vor Auslieferung
- Fragen/Probleme über Ticket-System („issues“) der Datenbank kommuniziert
- Qualitätssicherung erfolgt für jede PDF-Seite

Was machen wir? (2)



- nach Übergabe der Dateien auf den Bundestags-Server: Kontrolle der HTML-Blätterbücher auf Funktionalität in definierten Stichproben.
- außerdem: Durchsicht zu juristischen Aspekten (Urheber-, Persönlichkeits-, Geheimschutz-Rechte)



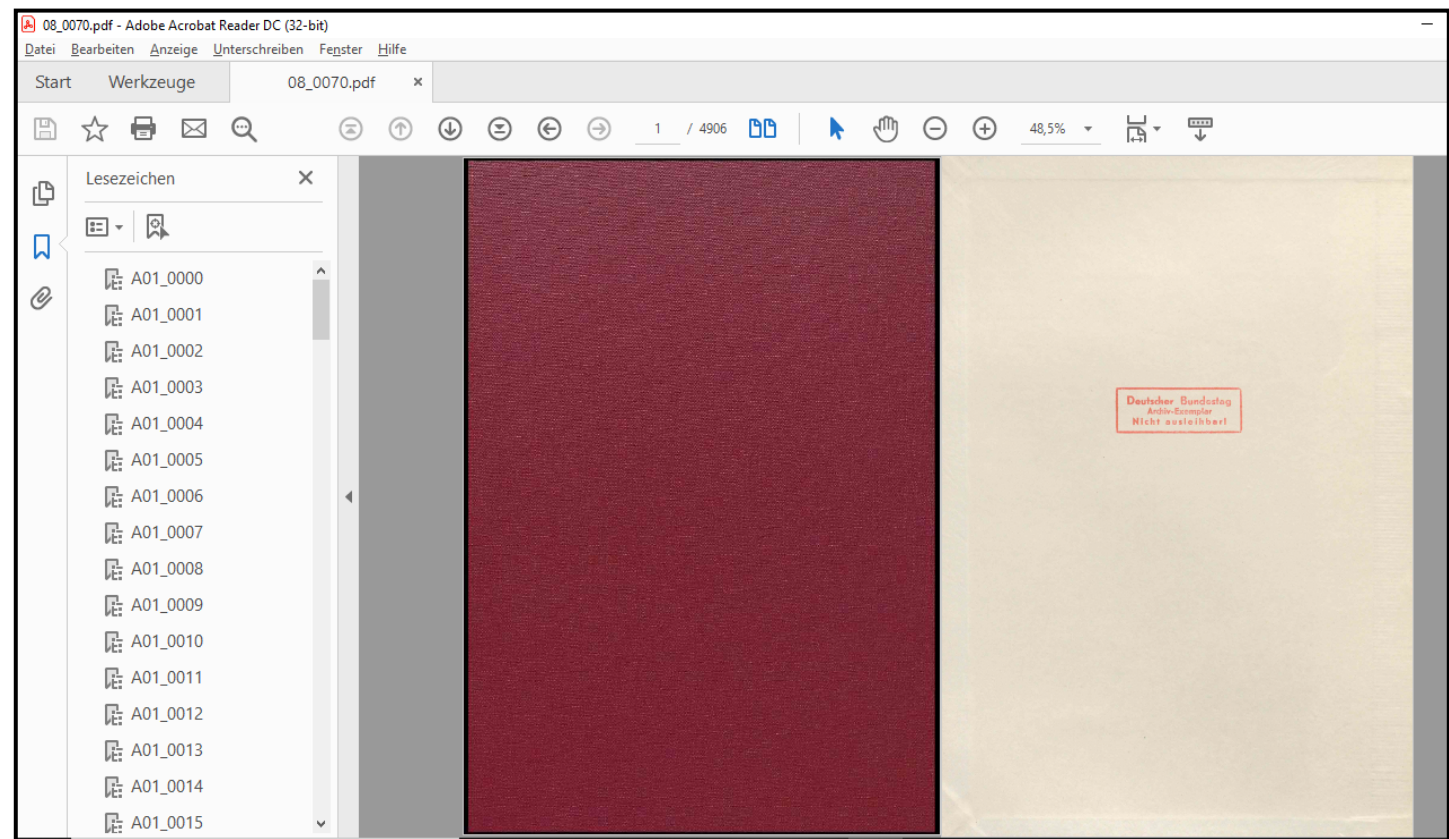
Wie sieht das Ergebnis aus?

- Farbtiefe: 24 Bit
- Auflösung: 300 dpi
- TIFFs auf LTO-Bändern
- weitere Dateien mit Sprungmarken im Inhaltsverzeichnis
 - PDF/A (für jede Dokumentation und für jedes Einzeldokument)
 - XML
 - HTML-Blätterbücher
- farbig, authentische Papierfarben und Formate (jedoch ggf. Drehung in Leserichtung)
- Austauschseiten bei leeren (Rück-) Seiten, Presseartikeln (Urheberrecht!), technisch bedingten Leerseiten
- OCR; bei schlechter Qualität der Originale: Darstellung dieser Bereiche als Image

Wie sieht das Ergebnis aus? (Beispiel PDF)

„Gesetz zur Änderung
der Strafprozessordnung“
vom 14.04.1978 (4900
Seiten, 6 Bände)

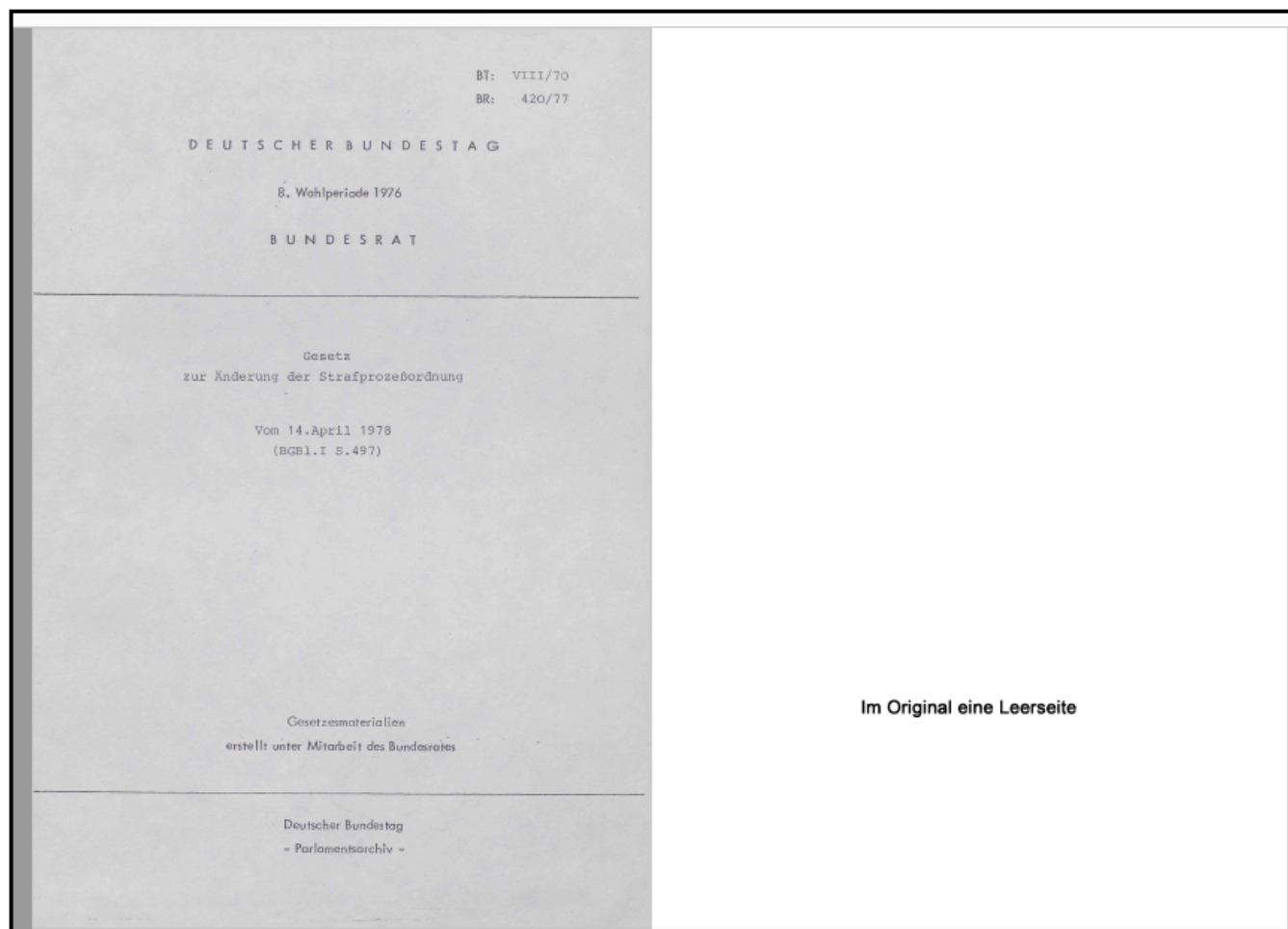
- links: Bookmarks als
Markierungen für
Einzeldokumente





Wie sieht das Ergebnis aus? (Beispiel PDF) (2)

- Titelblatt
- Rückseite als Austauschseite





Wie sieht das Ergebnis aus? (Beispiel PDF) (3)

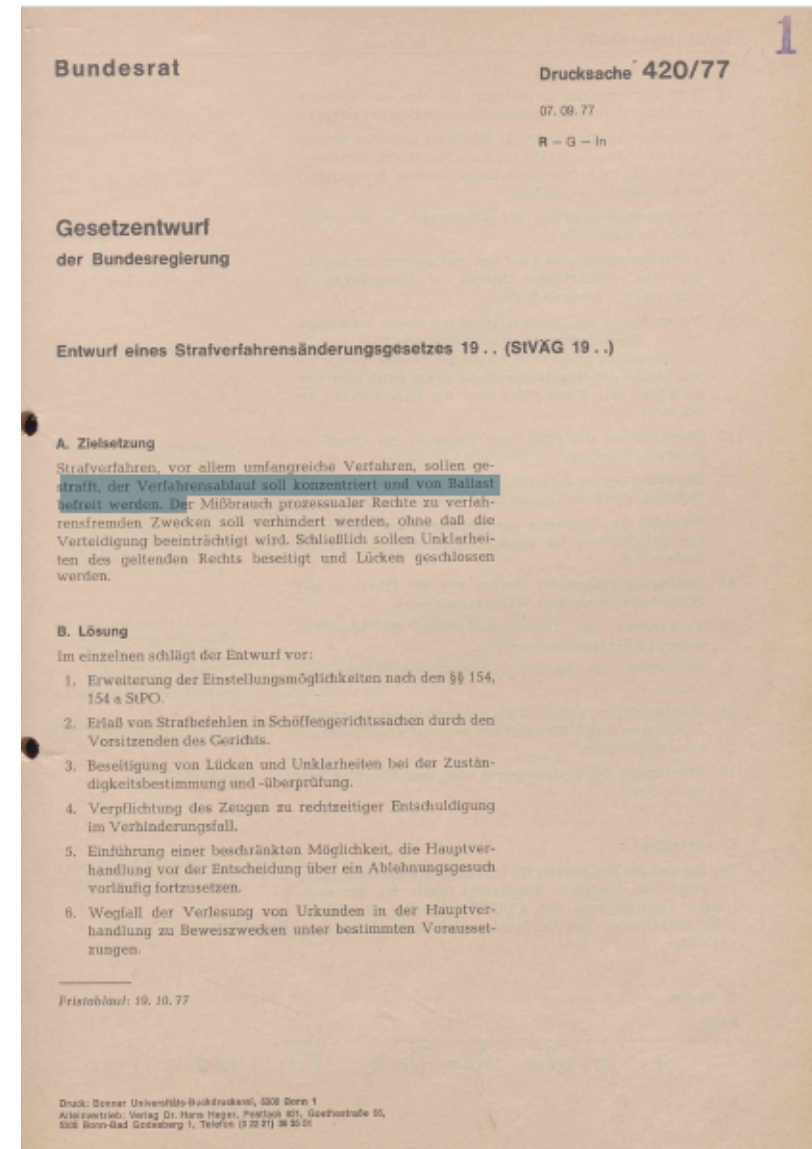
- Inhaltsverzeichnis
(Ausschnitt) mit
Markierung der Zone für
Mouseover- Bereich als
Link/Sprungmarke zum
Dokument

- 8 -

Lfd. Nr.	Datum	Drucksache/ Plenarsitz.	Körperschaft/ Verfasser	Titel / Dokumentbezeichnung Erläuterungen zum Inhalt
<u>Band A 1</u>				
Beratungen der Regierungsvorlage - BR-Drs.420/77 sowie des Gesetzesantrags - BR-Drs.535/77 im BR; Beratungen der Regierungsvorlage sowie der BT- Drs.8/322, 8/976, 8/1283 im BT (lfd.Nrn.1-33)				
B U N D E S R A T				
<u>1977</u>				
1	7.9.	420/77	BReg	<u>Gesetzentwurf</u> Entwurf eines Strafverfahrens- änderungsgesetzes 19... (StVKG 19....) Vorlage beim BR; AusschüB- zuweisung: R(f),G,In
<u>BR-Ausschußberatungen</u>				
2	21.9.		BR-In	<u>434.Sitz.</u> Beratungen und ÄndVorschl. zu <u>Art.1</u> <u>StPO</u> Nrn 55 11 138c nach 11 148 <u>Abgelehnte Antr.1</u>

Wie sieht das Ergebnis aus? (Beispiel PDF) (4)

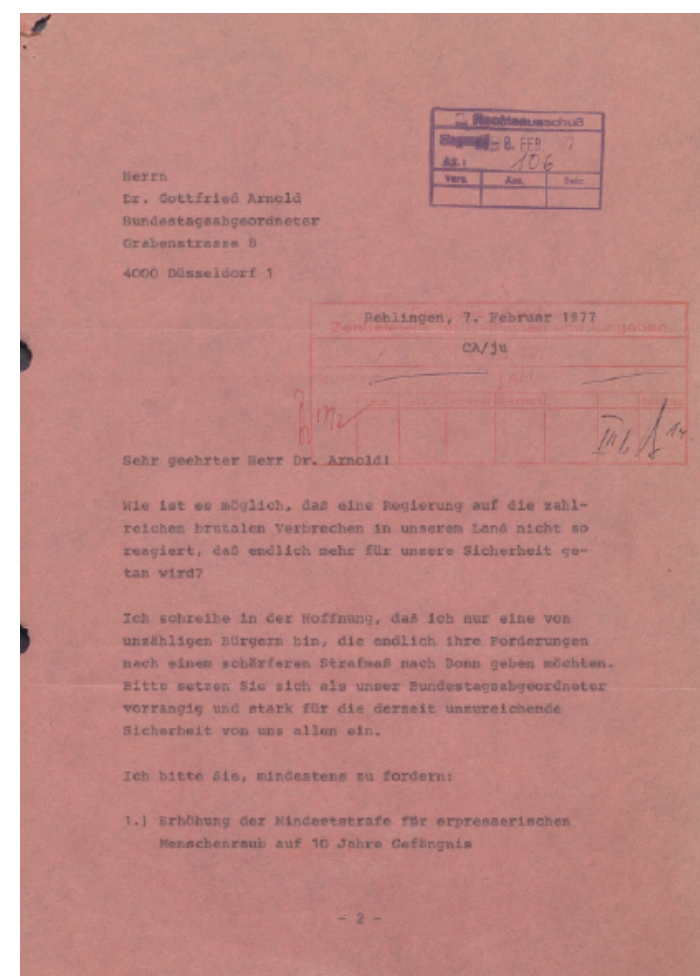
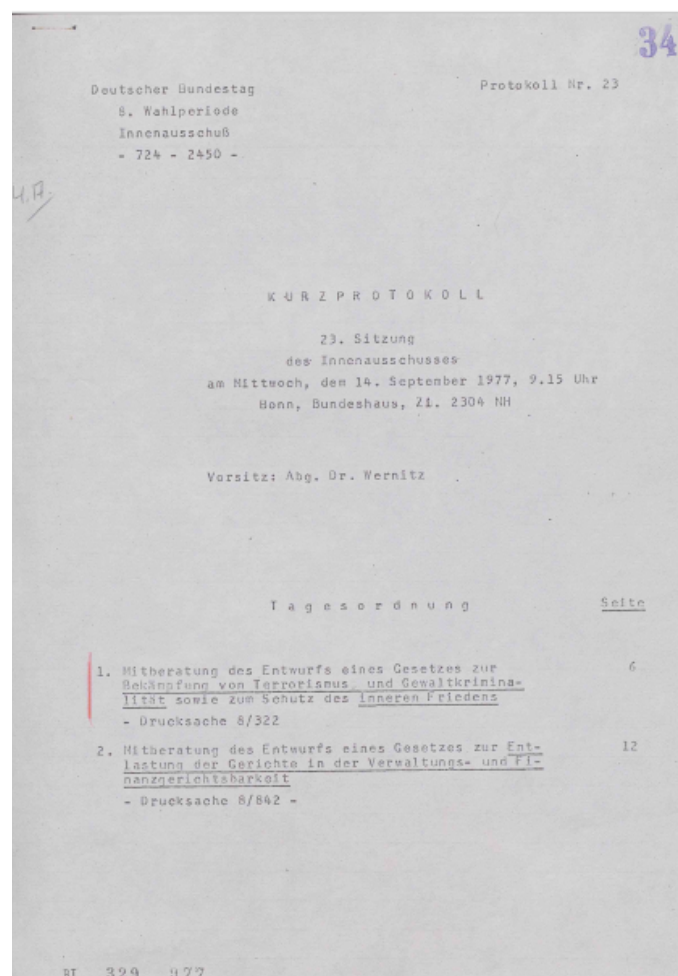
- Nach Auslösen des Links wird zum jeweiligen Dokument gesprungen.
- Markierung zeigt, dass Text erkannt wird, markierbar ist und nach Zeichenketten gesucht werden kann.





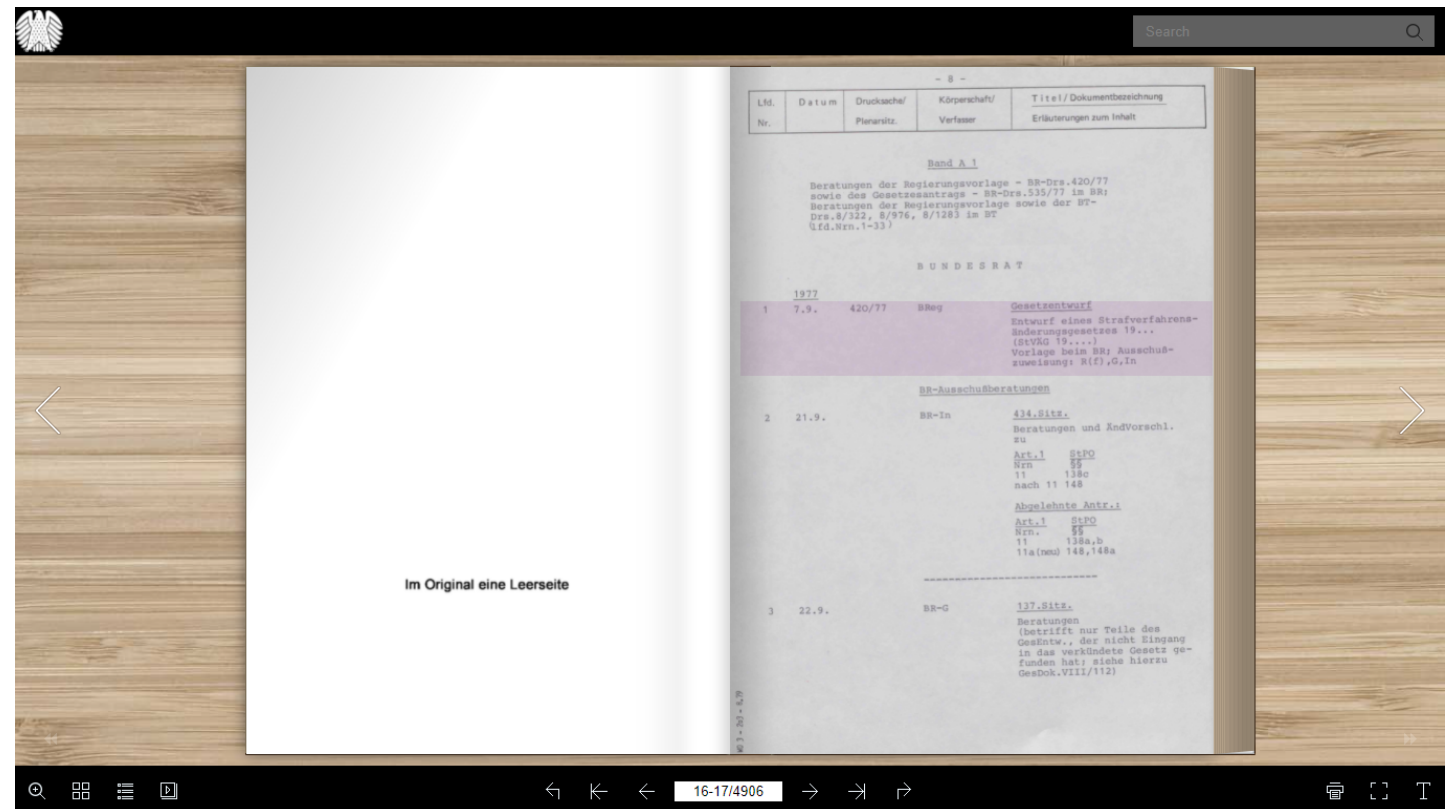
Wie sieht das Ergebnis aus? (Beispiel PDF) (5)

- Beispiele für weitere Dokumentenarten:
 - Ausschussprotokoll
 - Bürgerzuschrift
- gleichzeitig Beispiel für Schriftarten, Papierfarben, Stempel, ...



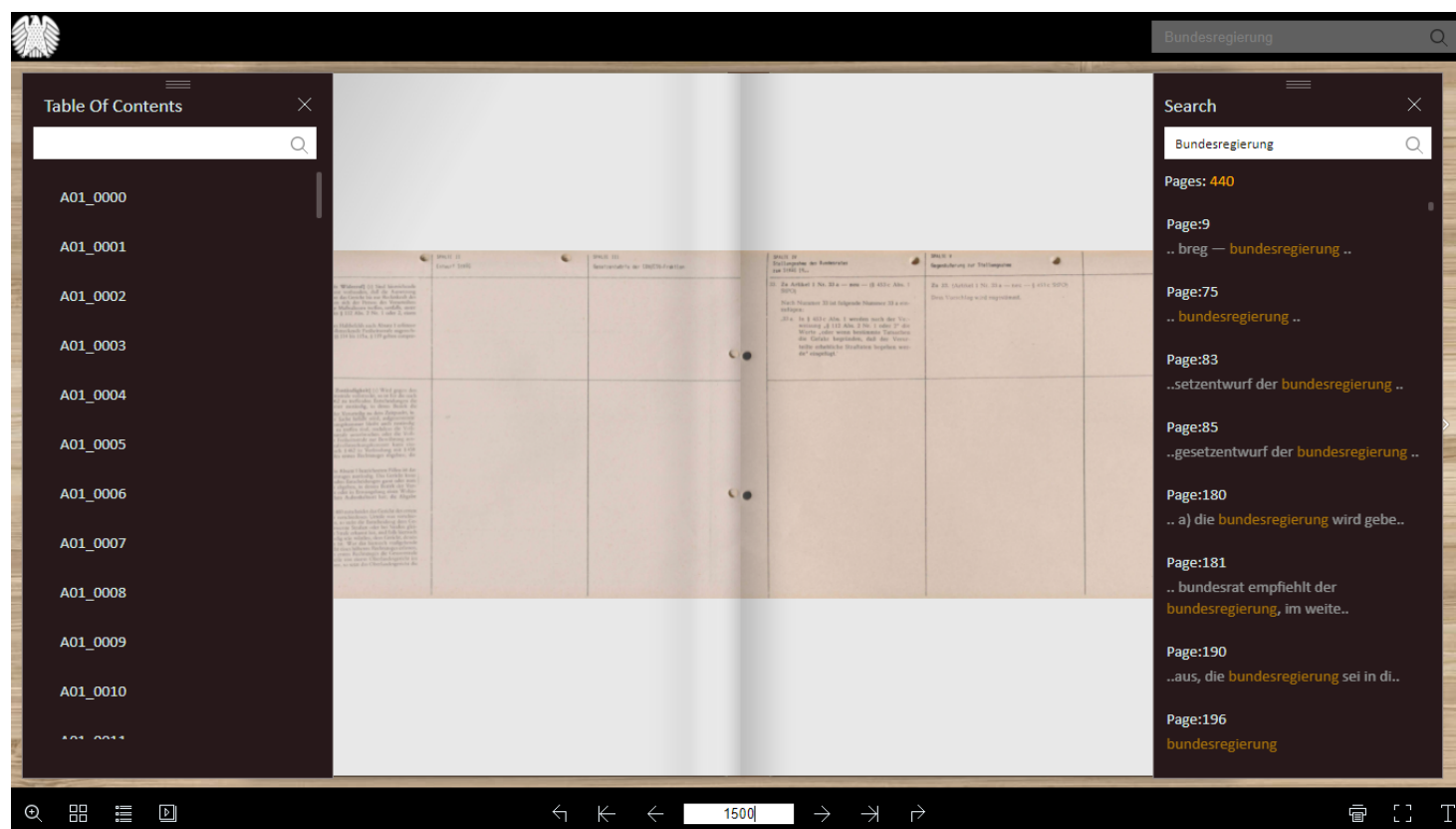
Wie sieht das Ergebnis aus? (Beispiel Flipbook)

- gleiches Dokument
- authentische Seitendarstellung (linke/rechte Seite)
- Inhaltsverzeichnis mit Markierung des Mouseover-Bereichs



Wie sieht das Ergebnis aus? (Beispiel Flipbook) (2)

- Flipbook-Features:
 - Begriffssuche mit Angabe der Trefferseiten (rechts)
 - Einzeldokumente als Bookmarks (links)



Worin liegen Stärken und Schwächen?

- Stärken
 - erheblicher Zeitgewinn durch Zugriff auf Workflow/Datenbank des Dienstleisters zur Qualitätssicherung, somit Fehlerbehebung größtenteils vor Auslieferung
 - äußerst zufriedenstellende Vertragsabwicklung (Qualität, Zeitrahmen, Rechnungslegung, Kommunikation, Transparenz...)
 - nachträgliche Änderungen auch von Flipbook-Dateien möglich
- Schwächen
 - hoher Aufwand zur Vorbereitung des Projektes in kurzer Zeit bis Kick off
 - Gesamtseitenzahl kann nur geschätzt werden (Kostenwagnis)
 - hoher Aufwand der eigenen QS (jedoch eigener Anspruch)
 - mäßige Sicherheitsstandards der Flipbook-Software

Fazit:

- Schwächen sind keine echten Schwachpunkte
- entscheidend: positive Vertragsabwicklung, v.a. Qualität

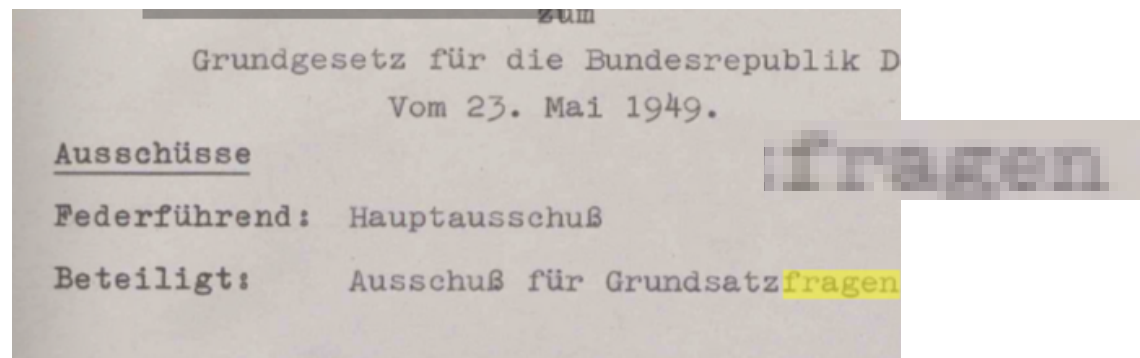


Was ist noch zu erwähnen?

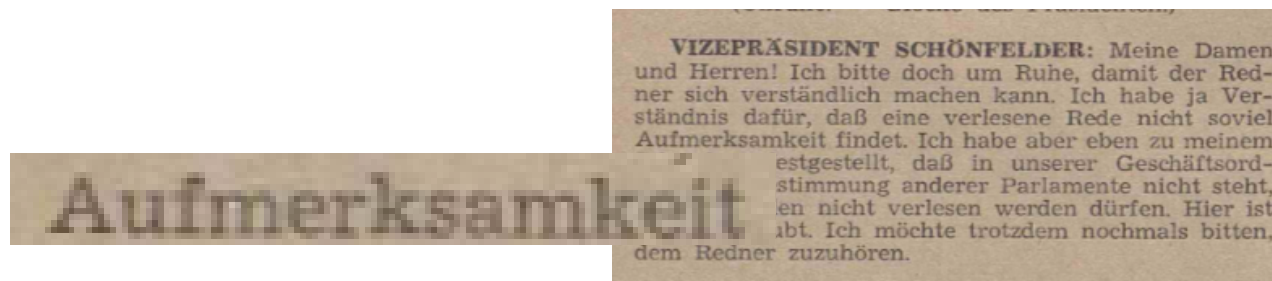
- Fokus: Bestandserhaltung
- Zugang/User-Interface: nicht im Mittelpunkt des Projektes; Nutzung auch weiterhin bei berechtigtem Interesse, entsprechend Anhang 2 der Geschäftsordnung des Deutschen Bundestages (Richtlinien für die Behandlung der Ausschussprotokolle und -drucksachen), da Ausschussunterlagen keine veröffentlichten Dokumente sind
- aktueller Stand (03/2022):
 - Flipbooks und PDFs der 10. WP wurden ausgeliefert
 - Bände der 11. und 12. WP werden aktuell gescannt und von uns durchgesehen
- Speicherbedarf/-prognose für 1.-16.WP:
 - PDFs: 5 TB
 - HTML-Flipbooks: 2 TB
- geplantes Projektende: 2. HJ 2023



Raum für Fragen



Herzlichen Dank für Ihre Aufmerksamkeit!



Kontakt: stefan.hauff-hartig@bundestag.de

Bildnachweis: Bundestagsadler:

Urheber: Prof. Ludwig Gies

Bearbeitung: 2008, Büro Uebele, Stuttgart

Folien 8 und 9:

iNCO Sp. z o.o.

alle anderen Folien:

Deutscher Bundestag / Stefan Hauff-Hartig