

# Archivierung von Websites mit dem DIMAG-Modul DIWI




Johannes Renz

**XArchiv? Automatisierung und Standardisierung bei der Digitalen  
Archivierung**

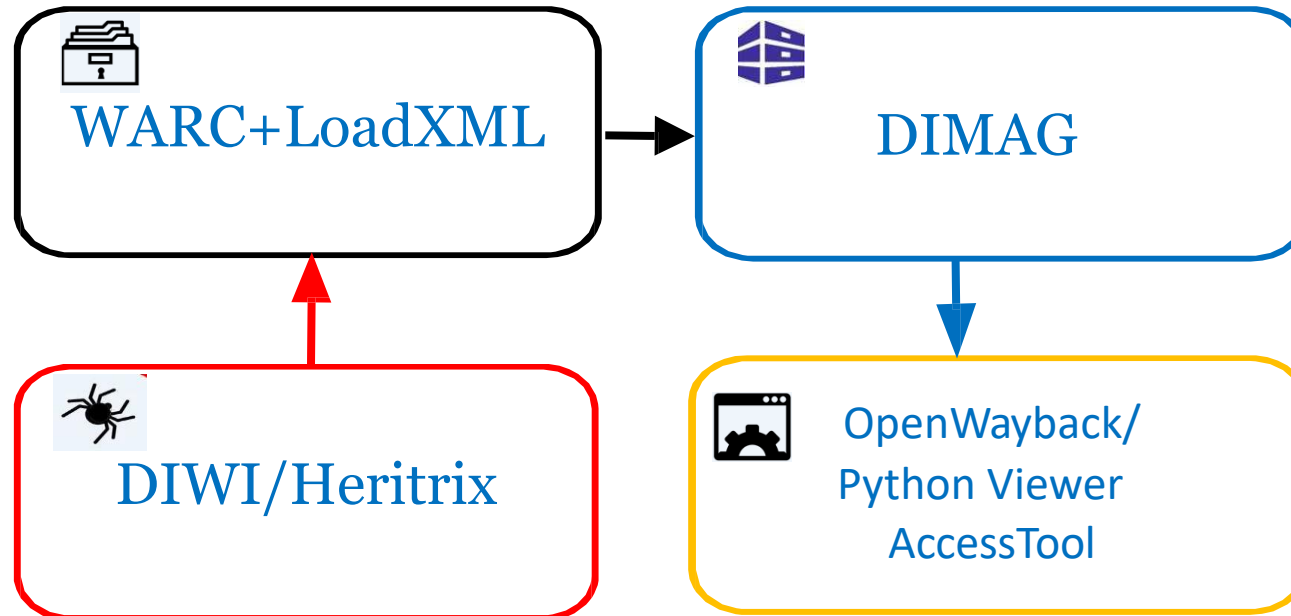
Online-Workshop der Generaldirektion der Staatlichen Archive Bayerns am  
27. Januar 2021

# Webarchivierung im Landesarchiv Baden-Württemberg

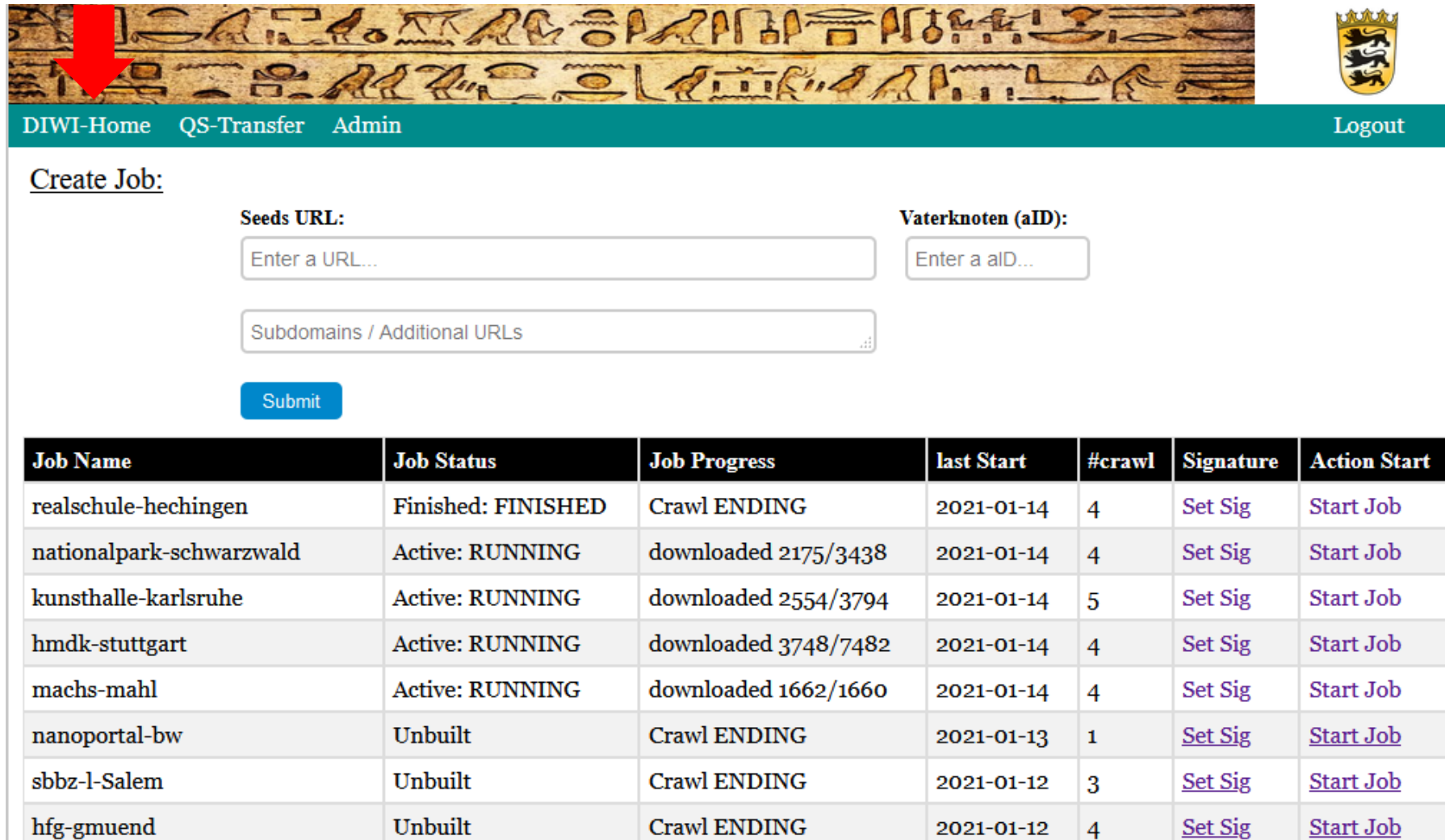


- ❖ Startschuss 2006 mit  HTTrack WEBSITE COPIER
- ❖ 12 Jahre auf separater Plattform bei öffentlichem Dienstleister
- ❖ Seit 2012 sukzessive Umstellung auf  HERIRIX (WARC) in Verbindung mit  Wayback Machine
- ❖ Startpaket: Websites u. a. der baden-württembergischen Landesministerien, ausgewählter Landesbehörden, (Fach-) Hochschulen und Kultureinrichtungen
- ❖ Ständige Evaluation und Erweiterung des Dokumentationsprofils
- ❖ Aktuell ca. 275 Webauftritte regelmäßig gecrawlt
- ❖ Mit Eventharvesting und offline geschalteten Sites: ca. 325 Webauftritte
- ❖ Seit 2018 Konzeption einer eigenen Webarchivierungslösung für den DIMAG-Verbund

# Konzeption von DIWI (DIMAG Website Ingest)



# DIWI-Home (Jobs anlegen und starten)



The screenshot shows the DIWI-Home interface. At the top, there is a navigation bar with links for 'DIWI-Home', 'QS-Transfer', 'Admin', and 'Logout'. Below the navigation bar, there is a 'Create Job:' section with a form. The form includes a 'Seeds URL:' field with a placeholder 'Enter a URL...', a 'Vaterknoten (aID):' field with a placeholder 'Enter a aID...', and a 'Subdomains / Additional URLs' field with a placeholder 'Subdomains / Additional URLs'. A blue 'Submit' button is located below the form. Below the form is a table with the following columns: 'Job Name', 'Job Status', 'Job Progress', 'last Start', '#crawl', 'Signature', and 'Action Start'. The table contains eight rows of job data.

Job Name	Job Status	Job Progress	last Start	#crawl	Signature	Action Start
realschule-hechingen	Finished: FINISHED	Crawl ENDING	2021-01-14	4	<a href="#">Set Sig</a>	<a href="#">Start Job</a>
nationalpark-schwarzwald	Active: RUNNING	downloaded 2175/3438	2021-01-14	4	<a href="#">Set Sig</a>	<a href="#">Start Job</a>
kunsthalle-karlsruhe	Active: RUNNING	downloaded 2554/3794	2021-01-14	5	<a href="#">Set Sig</a>	<a href="#">Start Job</a>
hmdk-stuttgart	Active: RUNNING	downloaded 3748/7482	2021-01-14	4	<a href="#">Set Sig</a>	<a href="#">Start Job</a>
machs-mahl	Active: RUNNING	downloaded 1662/1660	2021-01-14	4	<a href="#">Set Sig</a>	<a href="#">Start Job</a>
nanoportal-bw	Unbuilt	Crawl ENDING	2021-01-13	1	<a href="#">Set Sig</a>	<a href="#">Start Job</a>
sbbz-l-Salem	Unbuilt	Crawl ENDING	2021-01-12	3	<a href="#">Set Sig</a>	<a href="#">Start Job</a>
hfg-gmuend	Unbuilt	Crawl ENDING	2021-01-12	4	<a href="#">Set Sig</a>	<a href="#">Start Job</a>

# Evaluation und Transfer nach DIMAG



## Create Job

Seeds URL:

Vaterknoten (aID):

Job Name	Job Status	last Start	QS Way.	QS Py.	Action Transfer	Transfer Status
realschule-hechingen	Finished: FINISHED	2021-01-14	<a href="#">Viewer</a>	<a href="#">Viewer</a>	<a href="#">Transfer Job</a>	∅
nationalpark-schwarzwald	Active: RUNNING	2021-01-14	<a href="#">Viewer</a>	<a href="#">Viewer</a>	<a href="#">Transfer Job</a>	∅
kunsthalle-karlsruhe	Active: RUNNING	2021-01-14	<a href="#">Viewer</a>	<a href="#">Viewer</a>	<a href="#">Transfer Job</a>	∅
hmdk-stuttgart	Active: RUNNING	2021-01-14	<a href="#">Viewer</a>	<a href="#">Viewer</a>	<a href="#">Transfer Job</a>	∅
machs-mahl	Active: RUNNING	2021-01-14	<a href="#">Viewer</a>	<a href="#">Viewer</a>	<a href="#">Transfer Job</a>	∅
nanoportal-bw	Unbuilt	2021-01-13	<a href="#">Viewer</a>	<a href="#">Viewer</a>	<a href="#">Transfer Job</a>	✓
sbbz-l-Salem	Unbuilt	2021-01-12	<a href="#">Viewer</a>	<a href="#">Viewer</a>	<a href="#">Transfer Job</a>	✓
hfg-gmuend	Unbuilt	2021-01-12	<a href="#">Viewer</a>	<a href="#">Viewer</a>	<a href="#">Transfer Job</a>	✓

# ...archiviert!



Dokumentation **Strukturnavigation** Finden Sonstiges Scope

LABW  
Prod



Angemeldet als:

Johannes Renz

Passwort

ändern **Abmelden**

Rechte:

- Digitales Archiv (219-1) (Löschen)
- Hohenlohe-Zentralarchiv Neuenstein (3-1) (Löschen)
- Generallandesarchiv Karlsruhe (4-1) (Löschen)
- Staatsarchiv Wertheim (7-1) (Löschen)
- Dokumentation (D-1) (Löschen)
- Ministerien und zentrale Dienststellen seit 1945 (1-800) (Löschen)
- Landtag seit 1945 (1-2442358) (Löschen)

## Strukturnavigation

- Tektonik:
- D: Digitales Archiv
    - A: Archivalien
      - Hauptstaatsarchiv Stuttgart
        - Ministerien und zentrale Dienststellen seit...
          - Staatsministerium
            - Verwaltung, Recht, Finanzen
              - Staatsministerium: Websites (archivierte...
                - Homepage des Staatsministeriums...**

Gehe zu AID:

▲▼ Signatur / AID	▲▼ Titel	Details
Hier einfügen: (Gekoppelt mit AFIS)		
EA 1/190_DO 1	2006-11-14 <a href="http://www.stm.baden-wuerttemberg.de/">http://www.stm.baden-wuerttemberg.de/</a>	🔑
EA 1/190_DO 2	2007-01-15 <a href="http://www.stm.baden-wuerttemberg.de/">http://www.stm.baden-wuerttemberg.de/</a>	🔑
EA 1/190_DO 3	2007-03-14 <a href="http://www.stm.baden-wuerttemberg.de/">http://www.stm.baden-wuerttemberg.de/</a>	🔑
EA 1/190_DO 4	2007-05-15 <a href="http://www.stm.baden-wuerttemberg.de/">http://www.stm.baden-wuerttemberg.de/</a>	🔑
EA 1/190_DO 5	2007-07-16 <a href="http://www.stm.baden-wuerttemberg.de/">http://www.stm.baden-wuerttemberg.de/</a>	🔑
EA 1/190_DO 6	2007-09-19 <a href="http://www.stm.baden-wuerttemberg.de/">http://www.stm.baden-wuerttemberg.de/</a>	🔑
EA 1/190_DO 7	2007-11-14 <a href="http://www.stm.baden-wuerttemberg.de/">http://www.stm.baden-wuerttemberg.de/</a>	🔑
EA 1/190_DO 8	2008-01-14 <a href="http://www.stm.baden-wuerttemberg.de/">http://www.stm.baden-wuerttemberg.de/</a>	🔑
EA 1/190_DO 9	2008-03-17 <a href="http://www.stm.baden-wuerttemberg.de/">http://www.stm.baden-wuerttemberg.de/</a>	🔑
EA 1/190_DO 10	2008-05-14 <a href="http://www.stm.baden-wuerttemberg.de/">http://www.stm.baden-wuerttemberg.de/</a>	🔑
EA 1/190_DO 11	2008-11-14 <a href="http://www.stm.baden-wuerttemberg.de/">http://www.stm.baden-wuerttemberg.de/</a>	🔑
EA 1/190_DO 12	2009-05-14 <a href="http://www.stm.baden-wuerttemberg.de/">http://www.stm.baden-wuerttemberg.de/</a>	🔑
EA 1/190_DO 13	2009-11-16 <a href="http://www.stm.baden-wuerttemberg.de/">http://www.stm.baden-wuerttemberg.de/</a>	🔑
EA 1/190_DO 14	2010-05-26 <a href="http://www.stm.baden-wuerttemberg.de/">http://www.stm.baden-wuerttemberg.de/</a>	🔑
EA 1/190_DO 15	2010-11-15 <a href="http://www.stm.baden-wuerttemberg.de/">http://www.stm.baden-wuerttemberg.de/</a>	🔑
EA 1/190_DO 16	2011-03-22 <a href="http://www.stm.baden-wuerttemberg.de/">http://www.stm.baden-wuerttemberg.de/</a>	🔑
EA 1/190_DO 17	2011-05-16 <a href="http://www.stm.baden-wuerttemberg.de/">http://www.stm.baden-wuerttemberg.de/</a>	🔑
EA 1/190_DO 18	2011-06-16 <a href="http://www.stm.baden-wuerttemberg.de/">http://www.stm.baden-wuerttemberg.de/</a>	🔑
EA 1/190_DO 19	2011-11-15 <a href="http://www.stm.baden-wuerttemberg.de/">http://www.stm.baden-wuerttemberg.de/</a>	🔑
EA 1/190_DO 20	2011-11-29 <a href="http://www.stm.baden-wuerttemberg.de/">http://www.stm.baden-wuerttemberg.de/</a>	🔑
EA 1/190_DO 21	2012-04-27 <a href="http://www.stm.baden-wuerttemberg.de/">http://www.stm.baden-wuerttemberg.de/</a>	🔑

Websites 😊

Blogs 😊

Podcasts 😊

Videos 😊

Social Media 😊



**Vielen Dank!**

**Johannes Renz**

Referat Audiovisuelles und digitales Archivgut, digitale Fachaufgaben, Nutzung  
Landesarchiv Baden-Württemberg

Abteilung Hauptstaatsarchiv Stuttgart

Telefon: 0711/212-4330

E-Mail: [johannes.renz@la-bw.de](mailto:johannes.renz@la-bw.de)