
WISSENSBESTÄNDE IM NETZWERK DER SÄCHSISCHEN LANDES- ARBEITSGEMEINSCHAFT AUSEINANDERSETZUNG MIT DEM NATIONAL- SOZIALISMUS

Typisierung, Anforderungsanalyse und Digitalisierungskriterien der Bestände

| Konstantin Freybe |



SÄCHSISCHE
LANDESARBEITSGEMEINSCHAFT
Auseinandersetzung mit
dem Nationalsozialismus

Abstract

Kann digitale Infrastruktur der sLAG die erinnerungspolitische Arbeit stärken? Das Dokument präsentiert Ergebnisse einer Befragung aus dem Jahr 2023. Zwei Erkenntnisse aus dieser Befragung sind zentral: Heterogenität kennzeichnet das Netzwerk einerseits in Hinblick auf die Objekttypen in den Wissensbeständen, andererseits in der Ausrichtung der Erinnerungsarbeit. Digitale Werkzeuge kommen ebenfalls auf vielfältige Weise zum Einsatz, wobei deutlich wird, dass eine gemeinsame, das Netzwerk überstreckende Infrastruktur eine Entlastung darstellen würde. Hieraus leitet die Studie Anforderungen an Software und Metadatenstandards ab und betont deren Relevanz für die heterogenen Wissensbestände im kollaborativen sLAG-Netzwerk. Technische Expertise, wissenschaftliche Interoperabilität und verbindliche redaktionelle Workflows sind Schlüsselfaktoren. Diese Anforderungen werden schließlich anhand zweier Beispiele veranschaulicht, die aus dem Open Source Kontext der Apache Software Foundation stammen: Apache Solr und Jena.

Wissensbestände im Netzwerk der sächsischen Landesarbeitsgemeinschaft Auseinandersetzung mit dem Nationalsozialismus

Typisierung, Anforderungsanalyse und Digitalisierungskriterien der Bestände

	Abstract	2
1.	Einleitung und Fragestellungen des Gutachtens	4
2.	Befragung 2023	5
3.	Auswertung der Befragung	6
3.1	Befragung des sLAG-Netzwerks	6
3.2	Quantifizierende Auswertung	7
3.3	Anforderungen an die technische Ausstattung der sLAG	12
3.3.1	<i>Apache Software Foundation (ASF)</i>	14
3.3.2	<i>Apache Solr</i>	14
3.3.3	<i>Apache Jena</i>	15
3.4	(Meta-)Datenstandards digitaler und digitalisierter Wissensbestände	15
3.4.1	<i>FAIR- und CARE-Prinzipien</i>	16
3.4.2	<i>International Image Interoperability Framework</i>	17
3.4.3	<i>Research Description Framework</i>	18
3.4.4	<i>Metadata Encoding and Transmission Standard (METS), Metadata Object Description Schema (MODS)</i>	18
4.	Fazit: Anforderungen	19
A1	Integration und Auffindbarkeit heterogener Bestände	19
A1.1	<i>Zentrale Verwaltung eigener technischer Infrastruktur</i>	19
A1.2	<i>Flexible Datenmodellierung</i>	20
A2	Kollaboration und souveräne Kuration der Daten	20
A2.1	<i>Verbindliche redaktionelle Workflows für Erschließung und Kuration</i>	20
A2.2	<i>Betrieb der Datenbank auf Webserver</i>	20
5.	Zusammenfassung	21
	Anhang	22
	Impressum	23

1. Einleitung und Fragestellungen des Gutachtens

Die sächsische Landesarbeitsgemeinschaft Auseinandersetzung mit dem Nationalsozialismus (im Folgenden: sLAG) zeichnet sich durch eine partizipative Organisationsstruktur und ihren ausgeprägten Netzwerkcharakter aus. Mehr als 100 Vereine, Initiativen, Gedenkstätten und Einzelpersonen sind hier organisiert. Ihren Auftrag beschreibt die sLAG auf ihrer Website wie folgt:

„Wir sind ein geschichtspolitisches Netzwerk in Sachsen. Unsere Mitglieder arbeiten mit unterschiedlichen Zugängen und thematischen Schwerpunkten zur lokalen, regionalen und überregionalen NS-Geschichte.“ (sLAG 2024)

Vor diesem Hintergrund übernimmt die sLAG eine koordinierende, vernetzende und unterstützende Funktion innerhalb der geschichtspolitischen Erinnerungsarbeit in Sachsen. Gemeinsamer Bezugspunkt aller Mitglieder ist die Auseinandersetzung mit dem Nationalsozialismus in Sachsen. Zugleich verweist die Selbstdarstellung der sLAG auf die große Diversität innerhalb des Netzwerks – sowohl hinsichtlich thematischer Schwerpunkte als auch in Bezug auf methodische Zugänge zur Erinnerungsarbeit. Diese Vielfalt spiegelt sich zudem in den unterschiedlichen Organisationsformen und regionalen Verankerungen der Mitglieder wider.

Im Laufe der Zeit hat sich innerhalb des Netzwerks eine differenzierte Forschungs-, Sammlungs- und Arbeitskultur herausgebildet. Der kollektive Wissensbestand aller Mitglieder ist beachtlich: Er umfasst zahlreiche Quellen und Dokumente – oder deren Kopien – zur Erforschung der NS-Herrschaft in Sachsen. Derzeit sind diese Bestände jedoch institutionell fragmentiert und zumeist nur innerhalb der jeweiligen Vereine oder Initiativen zugänglich. Vor diesem Hintergrund denkt die sLAG darüber nach, ob und unter welchen Bedingungen eine stärkere gemeinsame Nutzung dieser Ressourcen die historische Forschung und die historisch-politische Bildungsarbeit im Netzwerk verbessern könnte. Dabei steht die Förderung ehrenamtlichen Engagements im Vordergrund, das hier – im Unterschied zur akademischen Forschung im engeren Sinne – als vorrangig betrachtet wird.

Erstens verspricht eine strukturierte Erschließung der vorhandenen Wissensbestände einen erheblichen Erkenntnisgewinn. Eine standardisierte Erfassung und Zugänglichkeit würde sowohl zivilgesellschaftlichen Initiativen als auch der akademischen Forschung ermöglichen, neue Perspektiven auf bekannte Themenfelder zu entwickeln. Statt isolierter Wissensinseln, die oft an Einzelpersonen gebunden und schwer zugänglich sind, sollte die Zukunft in der Verknüpfung und Zusammenarbeit über fachliche und regionale Grenzen hinweg liegen. Zweitens bestehen im Netzwerk – nicht zuletzt über die Fachstelle in Leipzig – bereits einschlägige fachliche Kompetenzen, die eine Orientierung an wissenschaftlichen Standards ermöglichen. In der Folge wäre ein einfacherer Wissenstransfer und Synergieeffekte zwischen ehrenamtlichen Initiativen, hauptamtlich geführten Gedenkstätten und akademischer Forschung zu erwarten.

Eine verstärkte Anbindung an wissenschaftlich-historische Forschung erscheint in diesem Zusammenhang nicht nur sinnvoll, sondern notwendig. Besonders relevant ist dabei die Anschlussfähig-

keit digitaler Ressourcen der sLAG an bestehende wissenschaftliche Infrastrukturen. Expertise in Bereichen wie der Provenienzforschung oder der Zugang zu institutionenübergreifenden Systemen – etwa der Nationalen Forschungsdateninfrastruktur (NFDI) oder europäischen Programmen wie Horizon Europa – könnten genutzt werden. Dies würde nicht nur den wissenschaftlichen Austausch fördern, sondern auch den Zugang zu Technologien zur nachhaltigen Speicherung digitaler Daten eröffnen.

Um die Anforderungen an eine solche Wissensinfrastruktur adäquat – das heißt unter Berücksichtigung der thematischen und institutionellen Vielfalt – formulieren zu können, wurde das vorliegende Gutachten von der sLAG gemeinsam mit dem Landesdigitalisierungsprogramm Sachsen 2023 in Auftrag gegeben. Im Zentrum stehen dabei folgende Fragen: Welche Mitglieder des Netzwerks verfügen über welche Typen von Wissensbeständen? Mit welchen digitalen Ressourcen arbeiten sie, und wie wird die damit verbundene Belastung wahrgenommen? Schließlich: Welche Anforderungen an eine gemeinsame digitale Wissensinfrastruktur lassen sich aus den Antworten auf diese Fragen ableiten?

2. Befragung 2023

In der Befragung wurde mit dem allgemein gefassten Begriff des „Wissensbestandes“ operiert. In diesem Kontext bezieht er sich auf einen Sammlungsbestand, wie er in Archiven und Bibliotheken üblich ist: textliche Originalquellen, Dokumentkopien, Fotografien und andere audiovisuelle Datenträger sowie Digitalisate der genannten Bestandstypen. Dies ermöglicht eine Offenheit im Ansprechen von Netzwerk-Mitgliedern, die nicht den Ausschluss bestimmter Bestandstypen schon zu Beginn impliziert. Unbeschadet dieser typologischen Offenheit ermöglicht Wissensbestand den Fokus auf die Erinnerungsarbeit im sLAG-Netzwerk. Im letzten Quartal 2023 wurden die Mitglieder der sLAG durch den Autoren des Gutachtens befragt, welche Bestände sie halten und wie sie sie für ihre Erinnerungsarbeit nutzen. Informationen über die vorhandenen Wissensbestände im Netzwerk sind entscheidend für die Entwicklung gemeinsamer digitaler Ressourcen, die für die Mitglieder zugänglich und an ihre Bedürfnisse angepasst werden können.

Das Nutzungsversprechen solcher geteilten Ressourcen liegt dabei in der Stärkung der Arbeit der sLAG und ihrer Mitglieder. Sie stellt in Aussicht, dass lokal agierende Mitglieder, die möglicherweise als Einzelpersonen oder kleine Organisationen auftreten, einerseits sichtbarer und andererseits weniger durch Aufbau und Pflege technisch aufwendiger, digitaler Ressourcen belastet werden. Aus landesweiter Perspektive besteht die Chance,

dass zuvor lokal oder regional begrenzt verfügbares Wissen nun überregional sichtbar, verfügbar und nutzbar wird. Einerseits kann so die Arbeit einzelner Mitglieder über deren räumliches Einzugsgebiet hinaus zur Kenntnis genommen werden. Andererseits ermöglichen sachangemessene digitale Ressourcen auch eine Zusammenschau von zuvor disparaten Wissensbeständen und damit neuen Erkenntnisse.

Ohne der eigentlichen Auswertung der Befragung zu weit vorgreifen zu wollen, ist es wichtig zu bedenken, dass die öffentliche Verfügbarkeit der Wissensbestände als langfristiges Ziel betrachtet werden sollte. Kurz- bis mittelfristig mangelt es der sLAG an einer Übersicht über die im Netzwerk existierenden Bestände. Deren Erschließung ist der erste Schritt. Eine Veröffentlichung der Bestände in digitalen Ressourcen kommt erst in Frage, wenn eigentums- und urheberrechtliche Fragen geklärt sind. Ein zentrales Findmittel für die netzwerkinterne Nutzung bildet somit die Voraussetzung für mögliche Veröffentlichungen.

An die Mitgliederbefragung selbst wurden mehrere Anforderungen gestellt. Sie durfte nicht zu stark vorgeben, was als Wissensbestand gilt und was nicht. Sie musste möglichst schnell und einfach zu beantworten sein, um die Mitglieder nicht über Gebühr von ihrer Arbeit abzuhalten. Eine Konzession an die zu erwartende Heterogenität der Wissensbestände im Netzwerk war der Verzicht auf Fragen, die auf die Quantifizierung von Beständen zielen. Stattdessen steht eine möglichst aussagekräftige Typisierung der vorhandenen Bestände im Zentrum, da sich dadurch – unter Berücksichtigung der zu erwartenden Heterogenität im Netzwerk – wichtige Anforderungen an digitale Wissensressourcen rekonstruieren lassen.

3. Auswertung der Befragung

In den folgenden Abschnitten stelle ich kurz den Aufbau der Befragung vor und gehe anschließend zur Darstellung der Ergebnisse über. Anschließend werde ich erklären, welche Anforderungen ich aus diesen Ergebnissen ableite. Die Diskussion der Anforderungen wird in zwei Teilen geschehen. Erstens werde ich Anforderungen an die technische Ausstattung vorstellen, mich dabei auf Software konzentrieren, Hardware hingegen vernachlässigen. Zweitens werde ich auf Leitlinien und Metadatenstandards zu sprechen kommen, deren Adaption eine hohe Anschlussfähigkeit an wissenschaftliches Datenmanagement und entsprechende Forschungsdateninfrastrukturen in Aussicht stellt.

3.1 Befragung des sLAG-Netzwerks

Die hier ausgewertete Befragung ist kein Werkzeug, das Zusammenhänge im Sinne quantitativer Sozialforschung ermitteln will, auch wenn quantifizierende Aussagen getroffen werden. Sie ist vielmehr ein teilstandardisiertes Werkzeug, das von Mitgliedern des sLAG-Netzwerks Auskünfte über ihre Erinnerungsarbeit und die dabei genutzten physischen wie digitalen Wissensressourcen erhebt. Das Ziel ist es, aus diesen Informationen wichtige Anforderungen an gemeinsame Ressourcen zu erkennen.

Technisch wurde die Befragung auf einer Unterseite der sLAG-Homepage umgesetzt. Der URL wurde anschließend per E-Mail an die Mitglieder des Netzwerks versendet, wie sie es vom internen E-Mail-Verteiler bereits kennen. Um die Mitglieder auf die Befragung einzustimmen, wurde eine Besprechung mit am Thema interessierten Mitgliedern der sLAG einberufen und im Rahmen des Landesweiten Arbeitstreffens 2023 – der jährlichen Mitgliedervollversammlung des Netzwerks – ein Kurzvortrag gehalten. Feedback aus beiden Veranstaltungen fand Eingang in das Design der Befragung.

Die Befragten wurden zu Beginn um persönliche Angaben gebeten. Aus datenschutzrechtlichen Gründen werden in diesem Dokument keine personenbezogenen Angaben reproduziert. Hierzu zählten Ort, Postleitzahl, E-Mail-Adresse, ob sie als Einzelperson oder Organisation und unter welchem Namen sie in Erscheinung treten. Die Unterscheidung zwischen Einzelperson und Organisation wird in den Ergebnissen zusammenfassend dargestellt.

Nach dieser Eröffnung begann der erste inhaltliche Block, der allgemein erfragte, ob Materialien für die Erinnerungsarbeit verwendet werden. Wer keine Materialien nutzt und z.B. ausschließlich mündlich und aus dem Gedächtnis vorträgt; ein theoretisches Szenario, dessen Möglichkeiten, diese Ressource zu teilen, sind denkbar eingeschränkt. Die nächste Frage führte durch den Bezug auf das Vorhandensein physischer Bestände („Ja“ oder „Nein“) die Unterscheidung physisch – digital ein. Die nächsten beiden Fragen waren auf physische Medientypen in diesen Beständen gerichtet und führten zwei Anwendungsszenarien in die Befragung ein: archivarisches und vermittelnde Nutzung. Hier wurden Antwortmöglichkeiten vorgegeben. Dies soll Aufschluss darüber geben, welche Anforderungen an welche Medientypen hinsichtlich ihrer Nutzung und Verfügbarkeit gestellt werden. Dieser thematische Block schloss mit einer Frage nach weiteren Medientypen und bot Gelegenheit,

auf Medientypen hinzuweisen, die in den vorigen Fragen nicht erwähnt wurden.

Der zweite thematische Block war auf digitale Ressourcen gerichtet. Einführend wurde abgefragt, ob die Mitglieder überhaupt digitale Ressourcen nutzen („Ja“ oder „Nein“). Anschließend und analog zum vorigen thematischen Block wurden Typen digitaler Ressourcen abgefragt, die von den Befragten genutzt werden. Die letzte Frage dieses Blocks galt einer Einschätzung des Aufwands für die Pflege digitaler Ressourcen, um festzustellen, ob eine geteilte Ressource als Entlastung in Frage kommt.

Abschließend wurde auch nach dem Aufwand für Lagerung und Bewahrung physischer Ressourcen gefragt. Hier wurde eine fünfstufige Skala angeboten, aus der Antworten ausgewählt werden konnten: „sehr gering“, „gering“, „neutral“, „hoch“ und „sehr hoch“. Außerdem wurde erfasst, wie lange die Befragten für die Beantwortung der Fragen gebraucht haben. Dies wird jedoch nicht in die Darstellung der Ergebnisse einfließen und dient vorrangig der Nachbereitung der Befragung. Die Befragung endete mit einem Dank an die Teilnehmenden.

Die Darstellung der Ergebnisse in den folgenden Abschnitten erfolgt unter Berücksichtigung der Persönlichkeitsrechte antwortender Personen. Entsprechend werden keine personenbezogenen Angaben reproduziert. Stattdessen werden aggregierte Ergebnisse vorgestellt.

3.2 Quantifizierende Auswertung

Insgesamt 36 gültig beantwortete Befragungen wurden im Erfassungszeitraum von September bis November 2023 registriert. Von diesen eingegangenen Antworten gaben 31 (86,11 %) an, als Organisation aufzutreten und 5 (13,89 %) als Einzelperson. Damit beteiligten sich 38 Prozent der sLAG-Mitglieder an der Umfrage.

Von 36 Antworten wiesen 27 (75 %) aus, die Befragten würden physische Ressourcen in ihrer Vermittlungsarbeit nutzen, 5 (13,89 %) verneinten diese Frage und 4 (11,11 %) machten dazu keine Angabe. 34 (94,4 %) der Befragten gaben hingegen an, dass sie digitale Ressourcen nutzen würden und nur zwei (5,6 %) verneinten.

Auf die Frage nach dem Aufwand zur Pflege und Bewahrung physischer Ressourcen sind nur zwei (5,6 %) Antworten verzeichnet. Eine Antwort lautet „neutral“, die andere „sehr gering“. Hinsichtlich des Aufwands für die Pflege digitaler Ressourcen fiel die Beteiligung deutlich höher aus. Abzüglich der zwei Antworten, die eine Nutzung digitaler Ressourcen verneinten, gaben 3 Befragte (8,82 %) den Aufwand als „sehr gering“, 4 (11,76 %) als „gering“, 12 (35,29 %) als „neutral“, 10 (29,41 %) als „hoch“ und 5 (14,71 %) als „sehr hoch“ an.

Aus den eingegangenen Antworten gingen die folgenden Häufigkeiten für physische Ressourcentypen hervor:

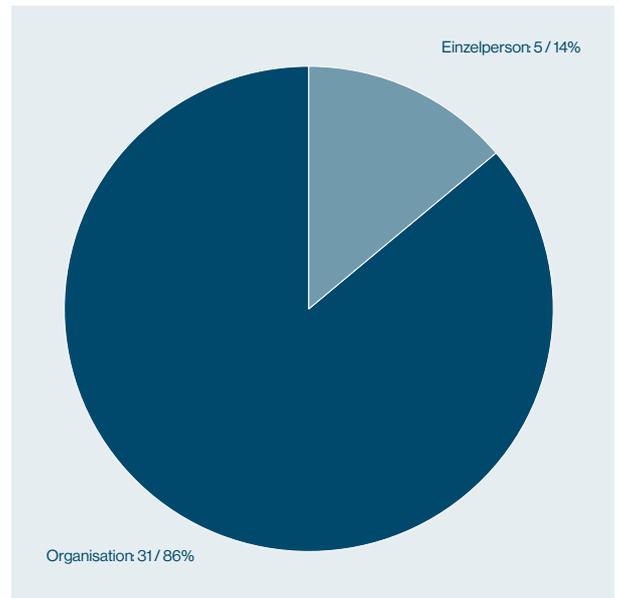


Abb: Verhältnis Organisation-Einzelperson

- Kopien von Akten, Briefen, Tagebüchern und/oder anderen Papieren: 24
- Fotografien: 24
- originale Akten, Briefe, Tagebücher und/oder andere Papiere: 19
- Kopien von historischen Karten: 12
- originale historische Karten: 6

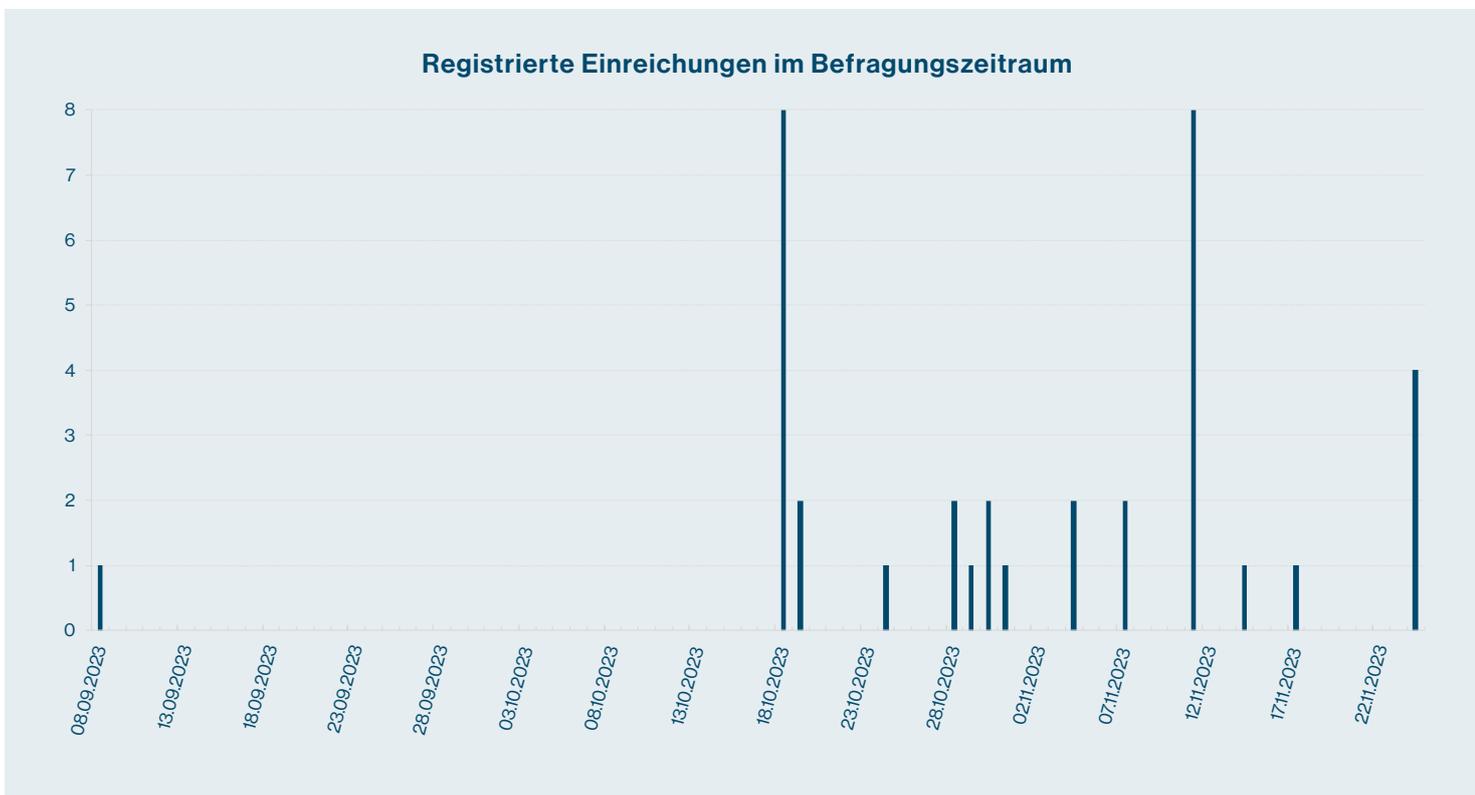


Abb: An 14 von 78 Tagen wurden insgesamt 36 Einreichungen registriert. (Erfassungszeitraum: September, Oktober und November 2023)

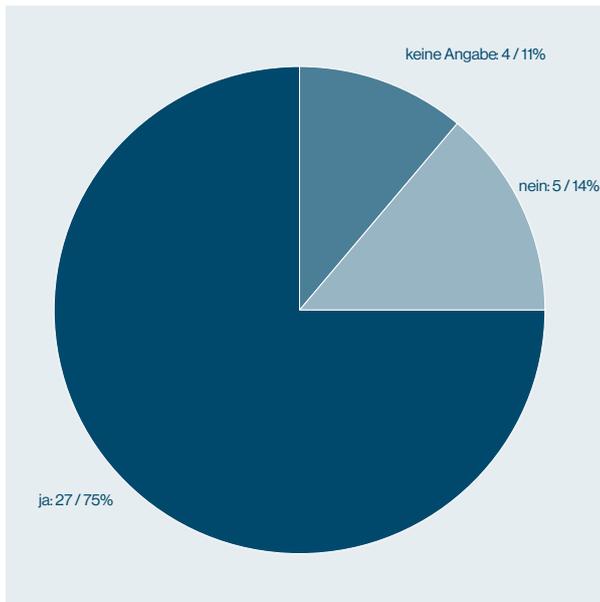


Abb: Angaben zur Nutzung physischer Ressourcen

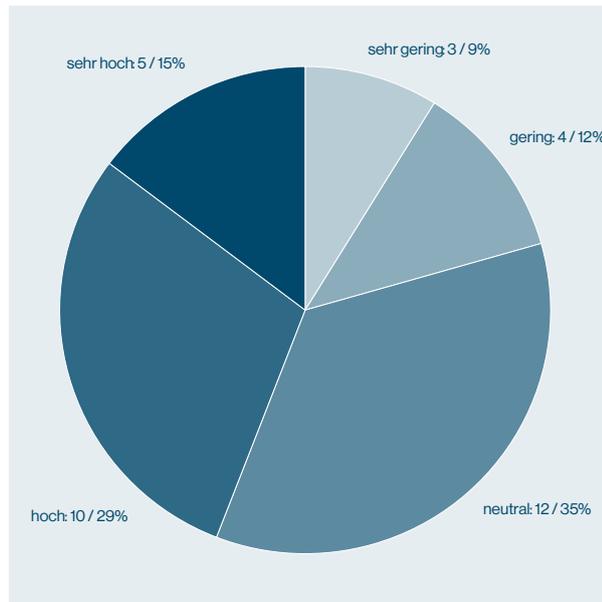


Abb: Pflegeaufwand für digitale Ressourcen

Hinsichtlich der Verwendung digitaler Ressourcen konnten folgende Häufigkeiten ermittelt werden:

- Festplatte in lokalem Computer: 25
- Webserven von Drittanbietern (z.B. Google Drive, Nextcloud ...): 19
- andere Trägermedien (z.B. CD-Rom, SD-Speicherkarte ...): 11
- eigener Webserver: 8
- Datenbanken: 1
- Eigener Netzwerkserv mit Datensicherung: 1
- Gedenkplätze.info: 1

Die Angabe von „Gedenkplätze.info“ scheint marginal, ist aber in der Einschätzung der vom sLAG-Netzwerk für die geschichtspolitische Arbeit mobilisierten Ressourcen keineswegs zu unterschätzen. Der URL <https://www.gedenkplaetze.info/> führt auf ein gemeinschaftliches Projekt eines Teils der Mitglieder des Netzwerks, das historisches Wissen über den Nationalsozialismus in Sachsen kontextualisiert und in Form einer interaktiven Karte georeferenzierte Orte mit Texten zur Geschichte dieser Orte verknüpft. Das Projekt ist 2018 mit dem Hans-Frankenthal-Preis der Stiftung-Auschwitz-Komitee ausgezeichnet worden. Die für das Projekt maßgeblich verantwortliche Organisation, der AKuBiZ e.V., wurde außerdem 2021 für seine

geschichtspolitische Arbeit auch unter Verweis auf dieses Projekt mit dem Obermayer-Award gewürdigt.

25 (69,44 %) der eingegangenen Antworten machten Gebrauch von der Option, in einem Freitext zusätzliche Angaben zu den in der Erinnerungsarbeit verwendeten Ressourcen zu machen. Diese Angaben erwiesen sich als aufschlussreich für die Diversität der Formen der Erinnerungsarbeit. Um diese Diversität darstellbar zu machen, wurden aus den Angaben sechs Typen rekonstruiert, die jedoch nicht vollständig trennscharf gegeneinander abgegrenzt werden können:

- „Literatur“ meint in diesem Kontext publizierte Texte in Form von Büchern oder Zeitschriften. Neun Antworten wurden diesem Typ zugeordnet.
- Die Bezeichnung „Öffentlichkeitsarbeit“ verweist auf all jene Materialien, die von den Befragten speziell für ihre Erinnerungsarbeit erzeugt wurden. Gästebücher oder Flugblätter sind anschauliche Beispiele für Materialien dieser Kategorie. Fünf Antworten wurden diesem Typ zugeordnet.
- Unter „Nachlass“ werden Objektbestände gefasst, die nach dem Versterben ihrer Eigentümer in den Besitz der Befragten übergegangen

sind. In einem Fall wird angegeben, dass die gesamte Wohnungseinrichtung einer Person verwahrt wird. Sieben Antworten wurden diesem Typ zugeordnet.

- „Archivgut“ ist insofern von Nachlässen verschieden, als es der weitere Begriff ist und auch formale Dokumente sowie solche mit diverserer Provenienz gefasst sind. Fünf Antworten wurden diesem Typ zugeordnet.
- Als „Forschungsdaten“ gelten an dieser Stelle Protokolle, Interviews, Notizen von Zeitzeug*innen. Vier Antworten wurden diesem Typ zugeordnet.
- Unter „Objektsammlung“ werden alle Sammlungen physischer Objekte gefasst, die so heterogen sind, dass sie sich einer näheren Zuordnung entziehen. Vier Antworten wurden diesem Typ zugeordnet.

Die nachstehende Grafik versucht die quantifizierende Darstellung der oben genannten Ressourcentypen. An dieser Stelle soll lediglich ein erster Überblick ermöglicht werden. Eine genauere Diskussion der in diesem Kontext gestellten Befunde wird im folgenden Abschnitt geleistet. Zwei Beobachtung sollen dem jedoch vorausgeschickt wer-

den:

1. Die Typen „Literatur“ und „Forschungsdaten“ verweisen auf vergleichsweise stark verbreitete Aufgabenstellungen, für deren Handhabung potenziell auf relativ lange Erfahrung zurückgegriffen werden kann. Das Verfügbarmachen von (Fach-)Literatur und Forschungsdaten deutet an, dass ein Teil des sLAG-Netzwerks nicht nur für die wissenschaftliche Forschung interessante Bestände hält, sondern dass im Netzwerk entsprechende fachliche Kompetenzen vorliegen, was wiederum ein Indiz für die Anschlussfähigkeit des Netzwerks an wissenschaftliche Forschung ist.
2. Das unter „Nachlass“ Verbuchte verweist stark auf Fragen der Provenienzforschung und sollte unbedingt ernst genommen werden. Auch wenn sich bereits abzeichnet, dass die Summe der Wissensbestände im sLAG-Netzwerk sich nicht mehrheitlich aus Nachlässen im hier verwendeten Sinne handelt, muss diesen Teilbeständen angemessen Rechnung bei der Formulierung von Anforderungen getragen werden, wenn sie innerhalb einer netzwerkübergreifenden digitalen Infrastruktur – und ggf. darüber hinaus – als Wissensressource

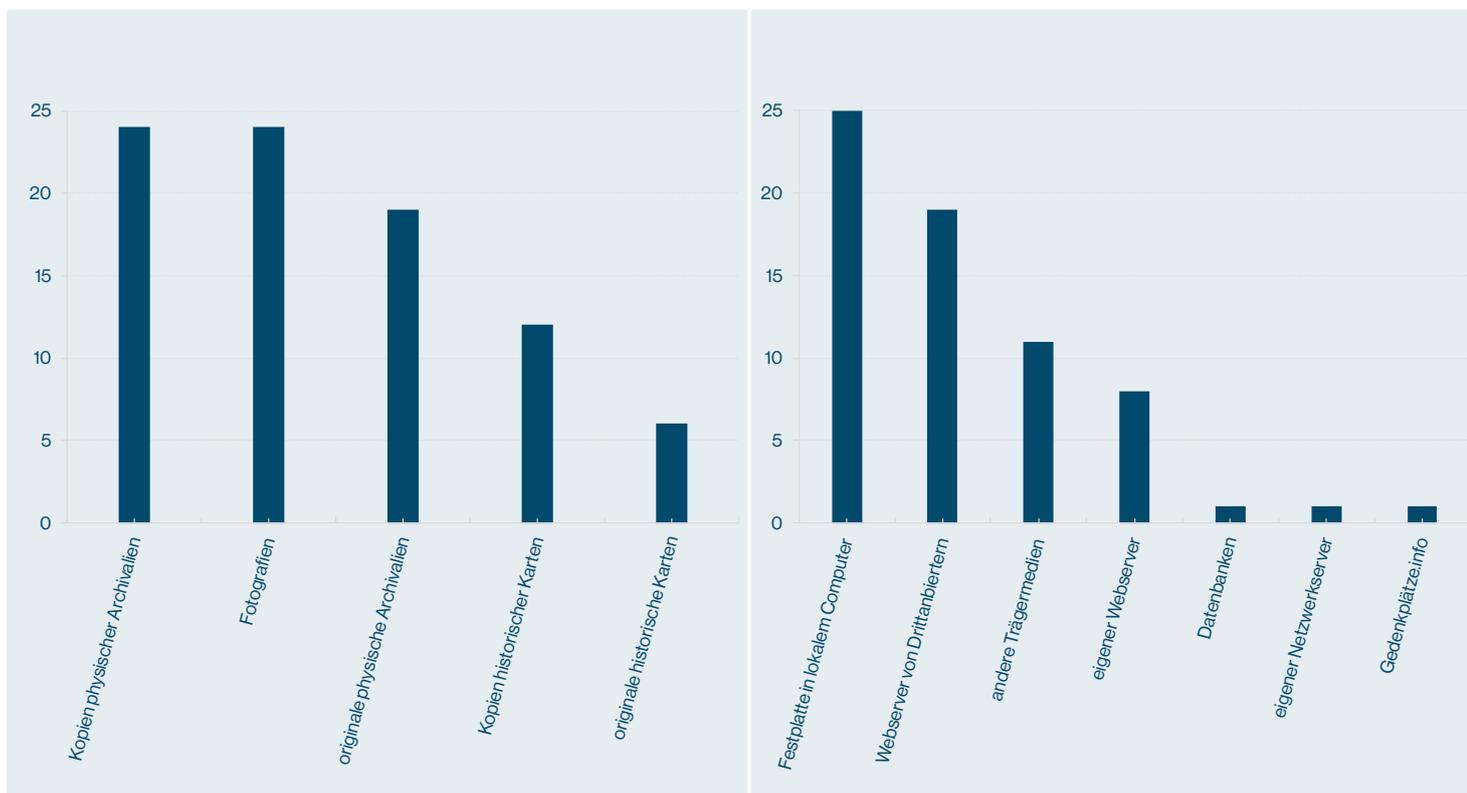


Abb: Häufigkeit physischer (links) und digitaler Ressourcentypen (rechts)

ertüchtigt werden sollen. Dies gilt besonders hinsichtlich rechtlicher Fragen.

Zusammenfassend ist die Situation des sLAG-Netzwerks auf mehreren Ebenen durch Heterogenität gekennzeichnet. Hinsichtlich der genutzten Ressourcentypen unterscheidet diese Studie zwei Register: digitale und physische Wissensbestände. Beide sind für die historisch-politische Bildungsarbeit im Netzwerk stark relevant. Beide sind außerdem stark heterogen. Auf der Seite physischer Wissensbestände herrscht über das gesamte Netzwerk betrachtet eine starke Materialdiversität. Hinsichtlich digitaler Wissensressourcen nutzen Mitglieder des Netzwerks verschiedene Infrastrukturkomponenten, die von Cloud-Speichern – sowohl als Service von Dritten als auch selbstständig betriebene – bis zu lokalen Speicherlösungen reichen. Letztere werden zu einem hohen Teil im Netzwerk für die Haltung und Verwaltung von Daten genutzt. Gleichzeitig sagen etwas mehr als 29% der registrierten Antworten aus, dass der Pflegeaufwand für digitale Ressourcen hoch sei. Die folgenden drei Punkten fassen die Ergebnisse grob zusammen:

1. *Materialdiversität und komplexe Erschließung:* Physische Wissensbestände im Netzwerk weisen eine hohe Materialdiversität auf. Ein beträchtlicher Teil dieser Bestände wurde als Archivgut charakterisiert, was Anlass zur Vermutung gibt, dass sich eine entsprechend systematische Erschließung lohnt, sofern sie nicht schon unternommen wurde.
2. *Keine geteilten, standardisierten digitalen Wissensressourcen:* Digitale Wissensbestände weisen keine einheitliche Struktur über das Netzwerk hinweg auf und die Wahl der technischen Infrastrukturkomponenten ist heterogen. Dass ein beträchtlicher Teil der Mitglieder des Netzwerks angibt, lokale Speicher für Haltung und Verwaltung digitaler Wissensressourcen einzusetzen, lässt sich als Anzeichen dafür lesen, dass die Praxis des Teilens technischer Infrastruktur noch nicht oder nur in geringem Maße

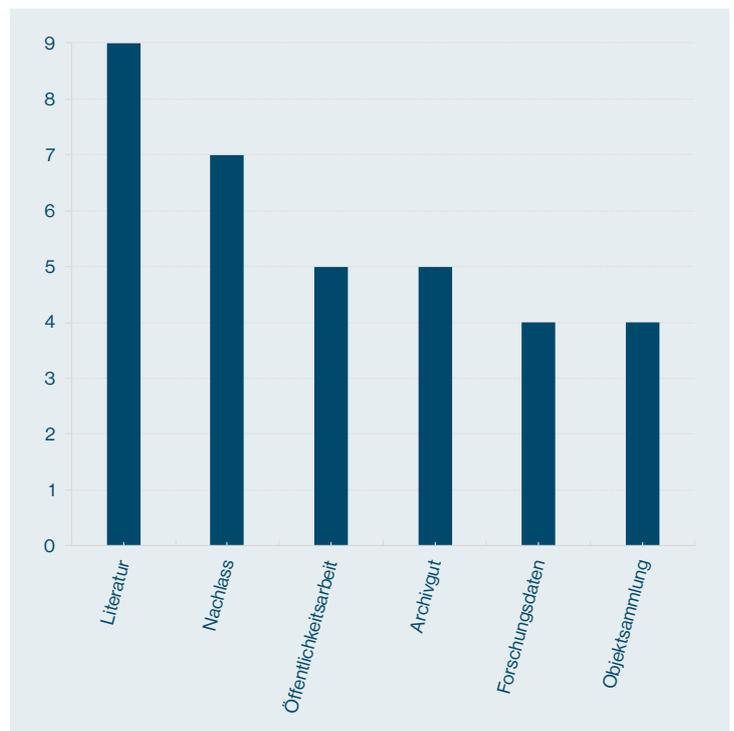


Abb: Andere nicht-digitale Materialien

etabliert ist. Dies mag auch der Heterogenität der Bestände und der Vielfalt der Zugänge zu Erinnerungsarbeit im Netzwerk geschuldet sein, aus denen spezifische und nicht generische Anforderungen folgen.

3. *Hoher Pflegeaufwand:* Dass ein beträchtlicher Teil der Mitglieder des Netzwerks angibt, die Pflege digitaler Ressourcen bringe einen hohen Aufwand mit sich, lässt sich so verstehen, dass hier eine Belastung empfunden wird und dieser Aufwand nicht oder nur unzureichend im Rahmen der eigenen Arbeit bewältigt werden kann. Folglich würde eine Entlastung in dieser Hinsicht eine Stärkung der Erinnerungsarbeit im Netzwerk bedeuten.

3.3 Anforderungen an die technische Ausstattung der sLAG

Der folgenden Diskussion zentraler Anforderungen an netzwerkübergreifende digitale Ressourcen möchte ich vorausschicken, dass die technische Ausstattung in diesem Zusammenhang auf Software und Metadatenstandards, nicht jedoch auf Hardware bezogen ist. Auf Grundlage der oben ausgebreiteten Befunde stellt das sLAG-Netzwerk zwei Anforderungen an die Digitalisierung der vorhandenen Wissensbestände: Souveränität und Gemeinschaftlichkeit.

1. *Souveränität*: Die Metadatenmodellierung und auch die Speicherung der primären Digitalisate sollte für eine souveräne Erschließung ausgerüstet sein. Mit souveräner Erschließung ist hier gemeint, dass das sLAG-Netzwerk eine eigene Infrastruktur für die Erfassung, Erschließung sowie die Kuration der erzeugten Bestandsdaten braucht. Dies zählt auf alle zentralen Ergebnisse der Befragung ein.

Souveräne (Meta-)Datenmodellierung ist dann eine Antwort auf die Materialvielfalt in den physischen Beständen, wenn sie den Mitgliedern genug Raum gibt, um die Spezifik ihrer physischen Artefakte in ihrer Erschließung zu berücksichtigen und gleichzeitig – im gesamten Netzwerk gedacht – geteilte und verbindliche Regeln für die Erschließung vorgibt, die ein Teilen der erzeugten (Meta-)Daten ermöglicht. Besonders der letzte Aspekt der Balance zwischen Materialspezifik und Standardisierung bzw. zwischen Offenheit und Verbindlichkeit verweist bereits auf die Praxis der gemeinsamen Nutzung digitaler Wissensressourcen und Infrastruktur.

Darüber hinaus sind Provenienzfragen noch klärungsbedürftig, weswegen für viele Ressourcen eine öffentliche Bereitstellung nicht ohne Weiteres in Frage kommt. Existierende Dienste, die im weiteren Sinne als Plattformen für das Erzeugen und Verwalten von geteilten digitalen Wissensressour-

cen auftreten, setzen voraus, dass rechtliche Fragen durch die Datengeber geklärt sind. Dies kann auf Basis des gegenwärtigen Erschließungsstandes nicht gewährleistet werden. Um mit einer geteilten Wissensressource die Erinnerungsarbeit im sLAG-Netzwerk stärken zu können und den Pflegeaufwand für einzelne Mitglieder zu senken, muss eine gemeinsame Infrastruktur also auf geteilten Erschließungsregeln aufsetzen und vom Netzwerk selbst souverän betrieben werden können.

2. *Gemeinschaftlichkeit*: Der Heterogenität der Wissensbestände im sLAG-Netzwerk ist geschuldet, dass die im Vorfeld dieses Gutachtens durchgeführte Befragung die Typisierung der Bestände höher priorisierte als deren Quantifizierung. Dem sLAG-Netzwerk ist dann mit einer geteilten digitalen Infrastruktur am meisten geholfen, wenn diese als übergreifendes Findmittel funktioniert, mit deren Hilfe schrittweise ermittelt werden kann, welche Bestandstypen wo, in welchem Umfang und mit welchem Rechtstatus vorliegen. Erst auf dieser Grundlage sind z.B. zielführende Aushandlungen mit Hinterbliebenen möglich, denn die Rekonstruktion von Provenienzen bedarf in diesem Fall gründlich strukturierter Vorarbeiten. Außerdem ermöglicht diese Vorgehensweise Synergien zwischen den Mitgliedern des Netzwerks. Bedenkt man die physischen Artefakte, die in ganz Sachsen verteilt bei Mitgliedern des Netzwerks eingelagert sind, wäre es für alle Beteiligten von Vorteil, über ein Recherchewerkzeug zu verfügen, mit denen Mitglieder im gesamten Netzwerk ermitteln könnten, welche Wissensbestände wo vorliegen und, z.B. anlässlich spezifischer Themen, Ressourcen zusammenziehen zu können.

Die Funktion der sLAG-Fachstelle in Leipzig als Netzwerkknoten legt nahe, dass Aufbau, Betrieb und Pflege digitaler Ressourcen, aber auch die Vermittlung von Standards und Workflows für Bestandserschließung, -digitalisierung und Metadatenmodellierung von ihr koordiniert und gesteuert werden. So kann die Fachstelle Sorge dafür tragen, dass dem Netzwerk eine zentrale Ressource

zur Verfügung steht, die eine souveräne Kuration der Inhalte ermöglicht. Aus der langfristigen Absicht öffentlicher Zugänglichkeit von Wissensressourcen ergibt sich wiederum eine weitere Aufgabe der Fachstelle: Die Vermittlung von redaktionellen Workflows an die Mitglieder. Ziel der Redaktionsarbeit sollte es sein, dass Mitglieder im Netzwerk ihre Expertise in die Bestandserschließung einbringen können, ohne jedoch mit Instandhaltung und Pflege technischer Komponenten belastet zu werden.

Die Suche nach geeigneten Softwarelösungen ist unter den obigen Gesichtspunkten also der Versuch die zentralen Anforderungen – Souveränität und Gemeinschaftlichkeit – in digitale Werkzeuge zu übersetzen. Doch auch abseits konkreter technischer Lösungen sind zwei Bereiche zu bedenken, die für eine erfolgreiche Etablierung einer geteilten digitalen Wissensressource im sLAG Netzwerk ebenfalls von Gewicht sind: redaktionelle Workflows und (Forschungs-)Datenmanagement.

Mit redaktionellen Workflows werden alle Maßnahmen beschrieben, die das Erfassen und Einpflegen neuer Daten, die Aktualisierung bestehender Daten (inklusive entsprechender Regeln) sowie deren Austausch und Bereitstellung betreffen. Datenmanagement verweist im Unterschied dazu auf diejenigen Maßnahmen, die auf langfristige Speicherung, nachhaltige Nutzbarkeit und die Absicherung der Daten vor Havariefällen zielen. Als Forschungsdatenmanagement sind diese Maßnahmen dann angemessen bezeichnet, wenn sie zusätzlich darauf zielen, die Wissensressourcen einer wissenschaftlichen Nutzung zugänglich zu machen und diese Zugänglichkeit langfristig sicherzustellen.

Souveränes Betreiben von Technologien und ihre gemeinschaftliche Nutzung greifen eng ineinander. Kann funktionierende Technologie nicht gewährleistet werden, steht auch keine digitale Ressource für die Erinnerungsarbeit zur Verfügung. Können Netzwerkmitglieder nicht mittels redaktioneller, standardisierter Workflows eine

hohe Datenqualität gewährleisten, steht es ähnlich schlecht. Sowohl Betrieb als auch Redaktionsarbeit bilden Aufgabenbereiche, die ggf. zusätzliches Personal erfordern, um verantwortungsvoll und gründlich bearbeitet zu werden. Eine konkrete Angabe in Personenarbeitsstunden ist mir zum gegenwärtigen Zeitpunkt nicht möglich, ich muss jedoch dringend darauf hinweisen, dass beides von enormer Wichtigkeit für das Gelingen eines solchen Infrastrukturprojektes ist und nicht *nebenbei* erledigt werden sollte.

Im Folgenden werde ich die Anforderungen an die technischen Komponenten anhand von zwei Vorschlägen für geeignete Software veranschaulichen. Beide stellen Alternativen zu *relationalen Datenbankstrukturen* dar. Eine relationale Datenbank speichert Daten in Form von Tabellen mit Zeilen und Spalten. Jede Tabelle hat eine feste Struktur mit definierten Feldern. Die Beziehungen zwischen den Tabellen werden durch Schlüssel hergestellt. Relationale Datenbanken sind zuverlässig und bieten Transaktionssicherheit, aber sie erlauben nur begrenzte Flexibilität bei der Datenmodellierung. SQL (kurz für: Structured Query Language) ist eine populäre und bewährte Abfragesprache für solche Datenbanken.

Eine Alternative dazu ist eine NoSQL-Datenbank (kurz für: not only SQL) wie das leistungsstarke *Apache Solr*, die eine flexible Datenmodellierung ermöglicht. Die Daten werden in JSON-ähnlichen (kurz für: Java Script Object Notation) Dokumenten gespeichert, ohne feste Schemata. Solr eignet sich hervorragend für Volltextsuche, Facettierung und Filterung von großen Datenmengen.

Eine weitere Alternative, die ich vorstellen möchte, ist eine Graphdatenbank namens *Apache Jena*, die Daten als Knoten und Beziehungen in einem Graphen speichert. Dieses flexible Datenmodell eignet sich besonders für die Darstellung komplexer, vernetzter Informationen. Jena unterstützt die Abfragesprache SPARQL und ermöglicht effiziente Traversierung und Analyse von Graphdaten. So-

wohl Apache Solr als auch Apache Jena bieten eine hohe Flexibilität bei der Datenmodellierung im Vergleich zu relationalen Datenbanken. Sie ermöglichen es, Daten in einer natürlicheren und anpassungsfähigeren Weise zu speichern und abzufragen. Beide sind Open Source Software und insofern für jedwede Nutzungsformen kostenfrei verfügbar. Sie sollen aber nur als konkrete Beispiele für zwei mögliche Technologien dienen. Insofern sind beide zu Zwecken der Veranschaulichung ausgewählt worden, um dem sLAG-Netzwerk bei der Auswahl entsprechender Technologien die Richtung weisen zu können.

Im Folgenden werde ich kurz die Apache Software Foundation vorstellen und anschließend die Projekte *Solr* und *Jena*.

3.3.1 Apache Software Foundation (ASF)

Die Apache Software Foundation entwickelt Software und formuliert in ihrem Mission Statement zwei zentrale Punkte: „The Apache Software Foundation (ASF) exists to provide software for the public good. We believe in the power of community over code.“ Gemäß der aktuellen Lizenz „Apache Licence, Version 2.0“ (Apache: Lizenz 2.0) bedeutet dies, dass die Nutzung, kommerziell oder nicht-kommerziell, von ASF Software solange frei erfolgt, wie keine Apache Markenzeichen verwendet und die begrenzte Garantie und Haftung von ASF anerkannt werden. Modifikationen am Code dürfen vorgenommen und auch an ASF zurückgespielt werden. Dies wird im Sinne von „the power of community over code“ als gemeinschaftlicher Beitrag zu den jeweiligen Softwareprojekten gewertet. Die folgende Darstellung dient lediglich dem Überblick und verzichtet auf die genaue Angabe von technischen Voraussetzungen.

3.3.2 Apache Solr

Apache Solr ist der Name einer Auskopplung von Apache Lucene und wurde 2021 zum „Top Level Projekt“ (Solr: Project) hochgestuft, was relative institutionelle Eigenständigkeit und eine verstärkte Allokation von Personalressourcen bedeutet. In

anderen Worten: Die ASF stuft dieses Projekt so wichtig ein, dass langfristige Pflege und Support der Software gewährleistet sind. Mit Solr bietet die ASF eine NoSQL Datenbank an, die „Advanced Full-Text Search Capabilities“ bereitstellt. Sie ist für große Datenmengen und ein hohes Aufkommen an online-Anfragen optimiert. Flexibles Konfigurieren einer Solr Instanz wird ebenso als Vorteil ausgewiesen wie die Verwendung mehrerer bewährter Standards wie XML, JSON und HTTP, die Erweiterbarkeit mit Plug-Ins oder auch ein eingebautes User-Interface für die Verwaltung der Solr-Instanz und den dort erzeugten Indizes. (Solr: Features)

Solr ist also geeignet für Datenmengen verschiedener Größenordnungen und kann im Bedarfsfall auch sehr hoch skalieren. Flexible Modellierung ermöglicht gegenstandsangemessene, d.h. der Vielfalt in den Wissensbeständen des sLAG-Netzwerks adäquate Beschreibung und Normierung. Letzteres ist wiederum von großem Wert für die Interoperabilität der erzeugten Daten, welche ihrerseits in geläufigen Dateiformaten exportiert und gemäß etablierter Standards (s.u.) modelliert werden können. Für die Integration bereits existierender digitaler Ressourcen in Solr stellt dies ebenfalls gute Voraussetzungen dar. Die konkrete Konversion bestehender Datenbanken in Solr-kompatible Formate hängt dabei jedoch von den konkret verwendeten Softwarekomponenten ab.

Die leistungsstarke Suche die Solr bietet und die Möglichkeit es auf Webservern zu betreiben, eröffnen zusätzlich eine Abfrage und Nutzung der Daten von externen Webapplikationen. Da ein graphisches Interface bereits mitgeliefert wird, entfällt beträchtlicher Gestaltungsaufwand für Nutzende. Auf der Solr-Projektwebsite werden weltweit agierende und stark frequentierte Dienste, die Solr nutzen, als Belege dafür angeführt, dass sich die Software auf hohem Niveau bewährt hat. Zu diesen Diensten zählen unter anderem: Bloomberg, Disney, DuckDuckGo, eBay, Instagram oder auch das Internet Archive. (Solr: Powered By)

3.3.3 Apache Jena

Apache Jena ist ein Java-basiertes Open-Source Framework für Linked Data und Semantic Web Anwendungen. Daten werden als Knoten konstruiert. Beziehungen zwischen Daten resp. Knoten werden als Kanten bezeichnet. Daten und ihre Beziehungen zu anderen werden als sogenannter Netzwerkgraph gespeichert, weswegen Datenbanken dieses Typs auch Graphdatenbank genannt werden. Das In-Beziehung-Setzen von Daten wird auch Verlinkung genannt, worauf die Bezeichnung Linked Data verweist. Eine weitere Bezeichnung für derartige Technologien lautet semantische Daten, weil die Beziehungen zwischen Daten nach dem Schema Subjekt-Prädikat-Objekt strukturiert sind. Als Subjekt gilt ein Datenknoten, das Prädikat beschreibt die Art der Beziehung näher und schließlich stellt das Objekt den Datenknoten dar, zu dem das Subjekt eine Beziehung unterhält.

Um Interoperabilität sicherzustellen, wurden Datenstandards vereinbart. „Resource Description Framework“ (w3: RDF) ist ein solcher Standard, mit dem auch Jena kompatibel ist. Insofern ist Jena ausgerüstet für das Verarbeiten von RDF/XML (Wikipedia: RDF/XML) und auch Turtle-Daten (Wikipedia: Turtle). Der Zugriff auf die Datenbank erfolgt in der Abfragesprache „SPARQL“. Mit „Fuseki“ stellt Apache Jena eine Infrastrukturkomponente bereit, die sowohl als Web-Applikation als auch als Standalone-Server für die Datenspeicherung eingesetzt werden kann. (Jena: Fuseki)

Falls die sLAG in Erwägung zieht, die eigenen Bestandsdaten bei angemessenem Erschließungsfortschritt als semantische Daten aufzubereiten und bereitzustellen, stellt ASF auch hierfür die Software bereit. Ein denkbare Szenario für die Nachnutzung solcher semantischen Daten im Kontext des sLAG-Netzwerks könnte z.B. im Teilen von Geodaten mit „gedenkplaetze.info“ liegen. Wir erinnern uns: diese Website wird zwar als eigenständiger Dienst betrieben und dient Nutzer*innen bei ihrer Erinnerungsarbeit gleichsam als Werk-

zeug. Das aus den Wissensbeständen des sLAG-Netzwerks erschlossene Wissen könnte dann gleichermaßen an einer zentralen Stelle gepflegt und, redaktionell überprüfte Georeferenzierungen vorausgesetzt, recht bald an andere Dienste weitergereicht und der politisch-historischen Bildungsarbeit bereitgestellt werden.

Die folgenden Abschnitte stellen kurz Verfahrensweisen und Standards der wissenschaftlichen Praxis vor. Dies erscheint mir mit Blick auf den Gegenstandsbereich und die Heterogenität der Wissensbestände zielführend, da diese Standards auch praktische Implikationen haben, sie z.B. für den Aufgabenbereich der Redaktionsarbeit (vgl. s.o.) von Bedeutung sind.

3.4 (Meta-)Datenstandards digitaler und digitalisierter Wissensbestände

Der bewusste Rückgriff auf (Meta-)Datenstandards aus dem Bereich wissenschaftlicher Forschung bietet meines Ermessens zwei Vorteile, die ich kurz hervorheben möchte: Einerseits steht auf diesem Weg in Aussicht, dass die Wissensressourcen der sLAG anschlussfähig an akademische Forschung werden und insofern auch Unterstützung aus entsprechenden Forschungscommunities erwartbar ist. Andererseits kann die Konformität mit etablierten wissenschaftlichen Standards als Verpflichtung auf gute wissenschaftliche Praxis im Sinne akademischer Gütekriterien kommuniziert werden, was die Erfolgchancen beim Einwerben von Fördermitteln aus der Wissenschaftsförderung potenziell steigert.

Neben der Metadatenmodellierung bildet die Digitalisierung physischer Wissensbestände einen wichtigen Meilenstein auf dem Weg zu einer netzwerkübergreifenden Ressource. Zwei Aspekte sind für das Erreichen dieses Meilensteins zentral. Der erste liegt in der Repräsentation des physischen Originals als digitale Datei. Medien wie z.B. Fotografien, Karten oder Schriftstücke im weiteren Sinne können via Scan oder Fotografie als digitale

Bilddateien repräsentiert werden. Andere Objekttypen wie Möbel, aber auch Kleidung oder Schmuckstücke im weiteren Sinne, können ebenfalls fotografiert werden. Ihre räumliche Qualität bliebe jedoch bei einer einzelnen Ansicht auf der Strecke, sodass hier entweder mehrere Bildaufnahmen aus unterschiedlichen Blickwinkeln oder z.B. 3D-Scans eine angemessenere Option darstellen. Audiovisuelle Teilbestände stellen wieder ganz eigene Anforderungen an ihre Digitalisierung. Damit solche Digitalisate in diesem auf die Repräsentation der Originale bezogenen Sinn kein abgeschiedenes Dasein in technischen Insellösungen fristen, ist die Konformität zu technischen Standards geboten.

Ein zweiter Aufgabenbereich liegt in der sachangemessenen Beschreibung der Medien bzw. Objekte der Wissensbestände. Die Modellierung solcher Metadaten ist entscheidend für die Auffindbarkeit und Bereitstellung von Digitalisaten. Um eine technische Anschlussfähigkeit an andere, ähnliche Sammlungen gewährleisten zu können, empfiehlt sich auch hier die Nutzung etablierter Standards. Wurde Metadatenmodellierung in den vorigen Abschnitten eher allgemein von den Möglichkeiten einer bestimmten Software her gedacht, stehen hier konkrete Metadatenstandards im Fokus, mit denen die Interoperabilität der Daten hergestellt werden soll.

Ein weiterer Aufgabenbereich liegt gewissermaßen quer zu den bisher umrissenen. Er richtet sich auf die Objekte in den Beständen. Die Frage, für welche Nutzer*innengruppen und Anwendungsfälle die Wissensbestände digitalisiert und bereitgestellt werden, ist hier ebenfalls wichtig. Denn bezieht man den stark partizipativen Charakter der Erinnerungsarbeit im sLAG-Netzwerk in diese Überlegungen mit ein, wird schnell deutlich, dass wissenschaftliche Nutzung zumindest kein exklu-

siv avisiertes Nutzungsszenario ist. Hiermit sind auf der einen Seite z.B. das Maß an Komplexität der Metadatenmodellierung oder Lizenzierungsfragen angesprochen. Auf der anderen Seite ist auch erheblich, ob die gewählte technische Umgebung der Arbeit der Mitglieder des sLAG-Netzwerks angemessen ist und eine Bereicherung darstellt. Hierbei sind nicht zuletzt Angaben zu technischer Ausstattung und zu Pflegeaufwand für digitale Ressourcen von Belang.

Abschließend ist wichtig, wie die Persistenz, also die dauerhafte Verfügbarkeit der Digitalisate, gewährleistet werden kann. Das betrifft einerseits die Art der Speicherung der Digitalisate. Andererseits ist unbedingt zu berücksichtigen, wie eine zuverlässige und dauerhafte Bereitstellung sichergestellt werden kann. Dieser Aspekt ist komplex. Bisher konnte als Anforderung der sLAG herausgearbeitet werden, dass sie über die Bestandsdaten souverän verfügen können muss. Insofern erscheint es sinnvoll, eine eigene Infrastruktur für die Datenspeicherung anzulegen. Um gegen technische oder menschliche Fehler und widerrechtliche externe Zugriffe bzw. Angriffe abgesichert zu sein, scheint eine Backup-Strategie zielführend, die auf externe Repositorien zurückgreift, in denen regelmäßig Sicherheitskopien gespeichert werden, die bei Bedarf zur Wiederherstellung genutzt werden können.

3.4.1 FAIR- und CARE-Prinzipien

Wie eingangs bereits angedeutet, erscheint eine Öffnung der sLAG-Ressourcen für wissenschaftliche und öffentliche Nutzungskontexte naheliegend. In den FAIR-Prinzipien (Findable, Resuable, Interoperable, Reusable) gibt es bereits etablierte Standards, um mit einem erweiterten Nutzer*innenkreis eigener Daten umzugehen. Gleichzeitig ist aber auch deutlich geworden, dass die sLAG und

ihre Mitglieder eine gewisse Autorität über ihre Bestände, deren digitale Repräsentation und schließlich auch über die eigenen digitalen Ressourcen behalten müssen. Dies wiederum nehme ich zum Anlass, eine Ergänzung zu den FAIR-Prinzipien vorzustellen: CARE (Collective Benefit, Authority of Control, Responsibility, Ethics).

Die FAIR-Prinzipien (FAIR 2024) bilden einen Leitfaden für die Gestaltung und Verwaltung von Forschungsdaten mit dem Ziel, deren Wiederverwendbarkeit und Zugänglichkeit zu verbessern. „Findable“ bezieht sich darauf, dass Daten leicht auffindbar sein sollten, sei es durch klare Kennzeichnung, eindeutige Identifikatoren oder Metadaten. „Accessible“ legt Wert darauf, dass Daten leicht zugänglich sein sollten, nicht nur für Maschinen, sondern auch für Menschen, und zwar unter Berücksichtigung von Datenschutz und Sicherheit. „Interoperable“ betont die Notwendigkeit, Daten in standardisierten Formaten bereitzustellen, um ihre Integration und Kombination zu erleichtern. „Reusable“ schließlich bedeutet, dass Daten in einer Weise bereitgestellt werden sollten, die ihre Wiederverwendung in verschiedenen Kontexten ermöglicht.

Dieser Satz an Richtlinien zielt also auf möglichst große Offenheit im Umgang mit Forschungsdaten. Im Vergleich dazu stellen die CARE-Prinzipien, wie die *Global Indigenous People Alliance* sie beispielsweise vorstellt (GIDA 2024), ein Plädoyer dafür dar, dass diese Öffnung nicht um jeden Preis und nicht ohne Rücksicht auf konkrete gesellschaftliche Verhältnisse durchgesetzt werden darf. Sie stellen gewissermaßen einen Versuch dar, Ansprüche an ein global gedachtes Gemeinwohl mit den Ansprüchen und Rechten indigener Völker auf Souveränität in Einklang zu bringen. Übertragen auf den Anwendungsbereich in dieser Studie und in Anbetracht

der Sensibilität und Spezifik der Wissensbestände im sLAG-Netzwerk erscheint es zielführend, den verantwortungsvollen Umgang mit (Forschungs-) Daten ebenfalls zu reflektieren.

Die CARE-Prinzipien stellen eine Ergänzung zu FAIR dar, indem sie soziale und ethische Überlegungen in den Umgang mit Daten einbeziehen. „Collective Benefit“ betont die Bedeutung von Datennutzung zum Gemeinwohl, während „Authority of Control“ auf die Verantwortung der Datenverwalter*innen hinsichtlich des Schutzes und der Verwendung der Daten hinweist. „Responsibility“ verweist darauf, dass alle Akteur*innen in der Datenkette für die Auswirkungen ihrer Handlungen verantwortlich sind. Schließlich legt „Ethics“ Wert auf die Beachtung ethischer Prinzipien bei der Datenerhebung, -verarbeitung und -nutzung. Durch die Integration von FAIR und CARE entsteht ein umfassender Ansatz für die verantwortungsbewusste und nachhaltige Handhabung von Forschungsdaten.

3.4.2 International Image Interoperability Framework

International Image Interoperability Framework (IIIF) wurde 2015 von einem Konsortium ins Leben gerufen, das aus Forschungs- und Nationalbibliotheken, Archiven, Software-Firmen und anderen besteht, die die folgenden Ziele gemeinsam verfolgen:

- „Providing researchers – and other audiences – an unprecedented level of uniform and rich access to digitized image, audio, and video resources hosted around the world
- Defining and maintaining a community-driven set of APIs that support interoperability between repositories
- Developing, cultivating, and documenting

shared technologies that provide a world-class user experience in viewing, comparing, manipulating, and annotating digitized visual and audio materials“

(IIIF: Consortium)

Zu den Gründungsmitgliedern dieses Konsortiums zählen unter anderen die Universität Oxford, die Britische Nationalbibliothek, die Universität Stanford, die Bayerische Staatsbibliothek sowie die Nationalbibliotheken Frankreichs und Norwegens. Aktuell zählt das Konsortium 67 Mitglieder. (IIIF: Mitglieder)

IIIF verfolgt die standardisierte Auslieferung von Bild- und audiovisuellen Dateien in verschiedene Umgebungen im Internet (vgl. IIIF: How IIIF Works). Der Anspruch von IIIF ist es, die basale Funktionalität der Darstellung dieser Dateien in Webbrowsern zu übersteigen:

„For images, that means enabling deep zoom, comparison, structure (...) and annotation. For audio/visual materials, that means being able to deliver complex structures (...) along with things like captions, transcriptions/translations, annotations, and more.“ (ebd.)

Es handelt sich bei IIIF also nicht nur um eine Metadatenstandardisierung, sondern zugleich auch um eine technische Umgebung, in der mehrere Programmierschnittstellen, Server und auch Präsentationskomponenten so zusammenarbeiten, dass ein hoher technischer Standard gewährleistet ist. Mit IIIF kann also die Zugänglichkeit und Nutzbarkeit digitaler Bildsammlungen im Internet ermöglicht werden. Die Verbreitung von IIIF in Kultureinrichtungen weltweit verspricht außerdem Anschlussfähigkeit an internationale Fach-Communities und bietet langfristig hohes Potenzial für Kollaborationen.

Die Dokumentation von IIIF ist für zwei Zielgruppen aufbereitet. Während Entwickler*innen eine detaillierte technische Dokumentation be-

reitgestellt wird (IIIF: Technische Doku), stehen Nutzer*innen ohne dezidierte Fähigkeiten der Softwareentwicklung Anleitungen und Trainingsressourcen zur Verfügung.

3.4.3 Research Description Framework

Das Research Description Framework (RDF) ist ein standardisiertes Modell zur Darstellung von Wissen im World Wide Web. Es dient dazu, Informationen strukturiert und maschinenlesbar zu verknüpfen, um so Daten effizient zu verarbeiten und auszutauschen. RDF basiert auf einer einfachen, aber leistungsstarken Dreiteilung von Aussagen, die aus Subjekt, Prädikat und Objekt bestehen, weswegen auch von semantischen Daten gesprochen wird. In solchen Tripel genannten Strukturen werden Daten so gespeichert, dass sie aufeinander verweisen und komplexe Wissensgraphen bilden.

In RDF wird Wissen in Form von Ressourcen und deren Beziehungen repräsentiert. Ressourcen werden durch Uniform Resource Identifiers (URI) referenziert. Beziehungen zwischen Ressourcen werden als Prädikate beschrieben. Diese semantische Struktur ermöglicht, dass nicht nur einzelne Informationen zu erfassen, sondern auch die Beziehungen zwischen solchen Informationen und Kontexten. RDF stellt gewissermaßen eine wichtige Grundlage für das Semantic Web dar, denn es stellt eine gemeinsame Sprache für den Austausch und die Integration von Wissen im Internet bereit. Die Verknüpfung von Datenquellen mittels RDF erleichtert die Suche, Analyse und Interpretation von Informationen.

3.4.4 Metadata Encoding and Transmission Standard (METS), Metadata Object Description Schema (MODS)

Im Bereich der Metadatenverwaltung für digitale Ressourcen sind der Metadata Encoding and Transmission Standard (METS) und das Metadata Object Description Schema (MODS) verbreitete Standards.¹ METS dient in diesem Zusammenhang als strukturiertes Schema zur Kodierung und Übertragung von Metadaten. Die sLAG tritt als Knoten

in einem Netzwerk von Einzelpersonen und Organisationen auf, die erinnerungspolitische Arbeit leisten. Um die Wissensbestände dieses Netzwerks zusammenführen zu können, ist ein gut strukturierter Austausch unverzichtbar. METS ist insbesondere für digitale Objekte mit komplexerer Struktur wie Bücher, Zeitschriften oder multimediale Inhalte entworfen worden. Dort wird die Struktur solcher Objekte definiert und die Integration verschiedener Metadatenstandards ermöglicht. Eine maschinenlesbare und plattformunabhängige Repräsentation der Metadaten wird dabei über XML sichergestellt.

Im Unterschied dazu ist MODS stärker auf bibliographische Metadaten gerichtet. Für bibliothekarische Gegenstandsbereiche stellt MODS eine Möglichkeit zur detaillierten Beschreibung von Ressourcen bereit. Für die Erinnerungsarbeit im sLAG-Netzwerk ist dies insbesondere von Belang, als dass erstens diese Arbeit auch auf Fachliteratur zurückgreift und zweitens z.B. als Ergebnis von Forschungsarbeiten in die Publikation von Fachliteratur mündet. Eine Verknüpfung von Wissensbeständen und Literatur, die auf sie Bezug nimmt, erscheint insofern naheliegend. METS und MODS bieten auch über dieses Beispiel hinaus eine gute Lösung für die Verwaltung und den Austausch von Metadaten. Für die Organisation und Pflege digitaler Sammlungen sind METS und MODS also gute Werkzeuge, mit denen die Strukturierung und Interoperabilität digitaler Sammlungen sinnvoll umgesetzt werden können.

¹ siehe Library of Congress

– METS: <https://www.loc.gov/standards/mets/>
– MODS: <https://www.loc.gov/standards/mods/>

4. Fazit: Anforderungen

A1. Integration und Auffindbarkeit heterogener Bestände

Die Auswertung der Befragung hat ergeben, dass die Mitglieder des sLAG-Netzwerks sehr verschiedene Objekttypen in ihren Sammlungen halten und auch hinsichtlich ihrer räumlichen Einzugsgebiete sowie in Bezug auf die Nutzung digitaler Ressourcen sehr verschieden sind. Daran knüpft sich der starke Verdacht, dass die Erschließungstiefe dieser Bestände ebenfalls äußerst unterschiedlich ist und besonders mit Blick auf Nachlässe oder urheberrechtlich relevante Dokumente Rechte- und Provenienzfragen dringend geklärt werden müssen. Hieraus ergeben sich weitere Anforderungen:

A1.1 Zentrale Verwaltung eigener technischer Infrastruktur

Die Möglichkeit rechtlich kontroverser oder zumindest noch klärungsbedürftiger Objekte in Wissensbeständen stellt ein beträchtliches Risiko für die Nutzbarkeit von Digitalisaten dar. Ein Einpflegen von digitalen Wissensbeständen in

existierende Dienste setzt rechtliche Klarheit voraus. Betreiber*innen solcher Dienste werden keine Verantwortung im Fall von Rechtsverletzungen übernehmen. Mit einer eigenen Infrastruktur kann die sLAG sicherstellen, dass die Erschließung der Wissensbestände im Netzwerk intern ungestört soweit voranschreiten kann, bis redaktionell die Veröffentlichung der Daten gerechtfertigt werden kann.

A1.2 Flexible Datenmodellierung

Aufgrund der Heterogenität der Wissensbestände ist eine flexible Datenmodellierung erforderlich, um ihre Zusammenführung in einer digitalen Ressource zu ermöglichen. Nur so kann ein lebensfähiger Kompromiss zwischen möglichst detaillierter und möglichst einheitlicher Beschreibung gelingen. Bei der Metadatenmodellierung sind Leitlinien wie FAIR und CARE sowie etablierte Metadatenstandards wie RDF, IIIF (besonders für Bilder und audiovisuelle Materialien) sowie METS/MODS (für bibliographische Ressourcen) hilfreich.

A2 Kollaboration und souveräne Kuration der Daten

Die Mitglieder des sLAG-Netzwerks verfügen über die Expertise für ihre Wissensbestände. Um diese Expertise in einer gemeinsamen digitalen Ressource fruchtbar zu machen, muss es ihnen ermöglicht werden, dass sie die Beschreibung ihrer Bestände in einer gemeinsamen Ressource zusammentragen. Außerdem erscheint es zielführend, redaktionelle Abläufe so einzurichten, dass Qualitätsmanagement ebenfalls kollaborativ erfolgt. Hieraus ergeben sich ebenfalls weitere Anforderungen:

A2.1 Verbindliche redaktionelle Workflows für Erschließung und Kuration

Die Einführung verbindlicher Workflows soll sicherstellen, dass eine hohe Datenqualität erreicht wird. Wenn dies nicht der Fall ist, sinkt die Nutzbarkeit der Daten drastisch und kann potenziell den Zweck einer gemeinsamen genutzten Ressource untergraben. Die Workflows sollen si-

cherstellen, dass Mitglieder in die Lage versetzt werden, selbständig ihre Bestände zu erfassen und dabei konform zu den vereinbarten Metadatenstandards vorzugehen.

A2.2 Betrieb der Datenbank auf Webserver

Aus technischer Sicht erfordert die kollaborative Nutzung einer gemeinsamen Datenbank im sLAG-Netzwerk, dass sie von Mitgliedern über das Internet erreicht werden kann. Daraus ergibt sich klar, dass die Datenbank auf Webservern betrieben werden muss. Aufgrund potenziell sensibler Inhalte muss der Zugang zur Datenbank beschränkbar sein. Außerdem sind Sicherheitsvorkehrungen unabdingbar. Basale Maßnahmen sind das Einrichten einer Nutzerkontenverwaltung und eine Backup-Strategie, die das Wiederherstellen vergangener Systemzustände erlaubt.

5. Zusammenfassung

Die sLAG ist eine wichtige Institution für das Erinnern an die nationalsozialistische Geschichte in Sachsen. Ihr Netzwerk ermöglicht Einzelpersonen und Organisationen unterschiedlicher Größe die Chance, ihre Forschungs- und Bildungsarbeit zu bündeln und zu unterstützen. Die Fachstelle der sLAG in Leipzig fungiert als Knotenpunkt in diesem Netzwerk und hat diese Studie veranlasst. Die Leitfrage lautet: Unter welchen Bedingungen kann digitale Infrastruktur die Arbeit der sLAG stärken und welche Voraussetzungen müssen für ihren Betrieb erfüllt sein?

Zusammenfassend lässt sich diese Frage so beantworten, dass digitale Dateninfrastruktur der sLAG dann am meisten nützt,

- wenn sie durch die sLAG und ihre Mitglieder selbst betrieben werden kann,
- wenn sie Derivate zu erzeugen erlaubt, die in geläufigen Dateiformaten wie XML oder JSON zum Austausch bereitgestellt werden können,
- wenn sie eine Datenmodellierung erlaubt, die flexibel genug ist, um die Balance zwischen einerseits detaillierter und andererseits einheitlicher und mindestens RDF-konformer Beschreibung der heterogenen Wissensbestände zu wahren,
- wenn sie kollaborativ genutzt wird, um die Wissensbestände tiefer zu erschließen und in

- der Folge zu einem robusten Findmittel für die sLAG-interne Nutzung heranwachsen kann,
- wenn auf ihrer Grundlage über Veröffentlichung und Nachnutzung der Wissensbestände entschieden werden kann.

Diese knappe Darstellung darf nicht über den personellen Aufwand hinwegtäuschen, den einerseits die Inbetriebnahme und Instandhaltung der technischen Infrastruktur mit sich bringt und der andererseits mit der Schulung der Akteur*innen im Netzwerk einhergeht. Sind diese Anforderungen erfüllt, bietet das sLAG-Netzwerk beträchtliches Potenzial die Geschichte des Nationalsozialismus in Sachsen heller auszuleuchten und auch unterschiedlichsten Gruppen zeitgemäß zugänglich zu machen.

Anhang

- sLAG 2024: <https://slag-aus-ns.de/>
- Apache – Lizenz 2.0: <https://www.apache.org/licenses/LICENSE-2.0>
- Solr – Project: <https://solr.apache.org/whoweare.html>
- Jena – Fuseki: <https://jena.apache.org/documentation/fuseki2/index.html>
- W3 – RDF: <https://www.w3.org/RDF/>

Wikipedia:

- NoSQL: <https://de.wikipedia.org/wiki/NoSQL>
- RDF/XML: <https://en.wikipedia.org/wiki/RDF/XML>
- Turtle: [https://en.wikipedia.org/wiki/Turtle_\(syntax\)](https://en.wikipedia.org/wiki/Turtle_(syntax))
- FAIR 2024: <https://www.go-fair.org/fair-principles/>
- GIDA 2024: <https://www.gida-global.org/care>

IIIF:

- Consortium: <https://iiif.io/community/consortium/>
- Mitglieder: <https://iiif.io/community/consortium/members/>
- How it Works: <https://iiif.io/get-started/how-iiif-works/>

Impressum

Autorenschaft:

Konstantin Freybe

Herausgeberschaft:

sächsische Landesarbeitsgemeinschaft
Auseinandersetzung mit dem Nationalsozialismus

Satz:

Jonas Kühne

Lektorat:

Jane Wegewitz

Redaktionsschluss:

19.07.2025

Spendenkonto: Förderverein der sLAG

IBAN: DE03 8605 5592 1090 2278 72

BIC: WELADE8LXXX

Sparkasse Leipzig

Betreff: Erinnern gestalten

gefördert durch:

Freistaat Sachsen

Landesdigitalisierungsprogramm für
Wissenschaft und Kultur



Diese Maßnahme wird mitfinanziert mit Steuermitteln auf Grundlage
des vom Sächsischen Landtag beschlossenen Haushaltes.

Sächsischen Landesarbeitsgemeinschaft
Auseinandersetzung mit dem Nationalsozialismus

Fachstelle NS-Erinnerungsarbeit
und Demokratiebildung
www.slag-aus-ns.de

2025