

GLES Open Science Strategie

Bucher, Hannah; Stroppe, Anne-Kathrin; Burger, Axel M.; Heycke, Tobias

Veröffentlichungsversion / Published Version

Arbeitspapier / working paper

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Bucher, H., Stroppe, A.-K., Burger, A. M., & Heycke, T. (2021). *GLES Open Science Strategie*. (GESIS Papers, 2021/03). Köln: GESIS - Leibniz-Institut für Sozialwissenschaften. <https://doi.org/10.21241/ssoar.68698>

Nutzungsbedingungen:

Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:

<https://creativecommons.org/licenses/by/4.0/deed.de>

Terms of use:

This document is made available under a CC BY Licence (Attribution). For more information see:

<https://creativecommons.org/licenses/by/4.0>

gesis

Leibniz-Institut
für Sozialwissenschaften

GESIS Papers

2021|03

German Longitudinal
Election Study



*Durchgeführt von der
Deutschen Gesellschaft für Wahlforschung und GESIS*

GLEES Open Science Strategie

*Hannah Bucher, Anne-Kathrin Stroppe,
Axel Burger & Tobias Heycke*

GESIS Papers 2021|03

GL ES Open Science Strategie

*Hannah Bucher, Anne-Kathrin Stroppe,
Axel Burger & Tobias Heycke*

GESIS Papers

GESIS – Leibniz-Institut für Sozialwissenschaften
Dauerbeobachtung der Gesellschaft
GESIS-Projektleitung German Longitudinal Election Study
Postfach 12 21 55
68072 Mannheim

E-Mail: gles@gesis.org

ISSN: 2364-3781 (Online)
Herausgeber,
Druck und Vertrieb: GESIS – Leibniz-Institut für Sozialwissenschaften
Unter Sachsenhausen 6-8, 50667 Köln

Das hier vorliegende Dokument beschreibt die strategische Ausrichtung der German Longitudinal Election Study (GLES) zur Förderung von Open Science (Offene Wissenschaft). Hierbei wird anhand von vier Grundpfeilern des Konzepts von Open Science – Open Methodology, Open Data, Open Source und Open Access – dargestellt, in welchen Bereichen die GLES derzeit gut abschneidet, in welchen Bereichen noch Verbesserungspotentiale vorhanden sind und welche Maßnahmen eingeleitet werden sollen, um die GLES nach Open Science Grundsätzen auszurichten. Die Umsetzung dieser Maßnahmen ist als langfristiger Prozess gedacht, bei dem existierende und neue Arbeitsprozesse sich am Ideal einer offenen Wissenschaftspraxis orientieren sollen. Ziel der hier vorgestellten Strategie soll es sein, sowohl Prozesse der Datenerhebung und Datenaufbereitung offen und transparent zu gestalten, als auch Forschende aktiv beim Praktizieren einer offenen Wissenschaft zu unterstützen.

Das Dokument wurde von den Autor*innen in enger Rücksprache mit dem gesamten GLES Team bei GESIS und der Koordinierungsgruppe (KG) der GLES verfasst. Besonderer Dank für Ihre Rückmeldungen und Unterstützung gehen darüberhinaus an Bernhard Miller und Pascal Siegers.

Inhalt

1	Hintergrund.....	5
1.1	Open Science als Chance für die GLES.....	6
2	Open Methodology	7
2.1	Status Quo	8
2.2	Verbesserungspotenzial	8
2.3	Maßnahmen	9
2.3.1	Publikationen zu spezifischen Inhalten des GLES Fragenprogramms	9
2.3.2	Publikation der Codes von Analysen	9
2.3.3	Veröffentlichung des Frageprogramms der GLES mit dem Feldstart	9
2.3.4	Veröffentlichung von englischen Übersetzungen des Fragenprogramms der GLES mit dem Feldstart	10
2.3.5	GLES Präregistrierungs-Challenge	10
3	Open Data.....	11
3.1	Status Quo	11
3.2	Verbesserungspotenzial	11
3.3	Maßnahmen	12
3.3.1	Veröffentlichung von englischen Übersetzungen der Datensätze.....	12
4	Open Access	13
4.1	Status Quo	13
4.2	Verbesserungspotenzial	13
4.3	Maßnahmen	13
4.3.1	Open Access Publikationsstrategie.....	13
4.3.2	Hervorhebung von Open Access Publikationen mit GLES Daten.....	13
4.3.3	Gewährleistung der Barrierefreiheit von GLES Publikationen	13
5	Open Source.....	14
5.1	Status Quo	14
5.2	Verbesserungspotenzial	14
5.3	Maßnahmen	14
5.3.1	Bereitstellung der Daten im CSV-Dateiformat.....	14
	References	15

1 Hintergrund

„Open Science is a system change allowing for better science through open and collaborative ways of producing and sharing knowledge and data, as early as possible in the research process, and for communicating and sharing results. This new approach affects research institutions and science practices by bringing about new ways of funding, evaluating and rewarding researchers. Open Science increases the quality and impact of science by fostering reproducibility and interdisciplinarity. It makes science more efficient through better sharing of resources, more reliable through better verification and more responsive to society’s needs“ (European Union, 2019).

Die Idee der offenen Wissenschaft (Open Science) ist es, den Prozess wissenschaftlicher Erkenntnisgewinnung offen, transparent und kooperativ auszurichten. Vor dem Hintergrund, dass durch eine transparente Wissenschaftspraxis der Grundstein für Widerlegbarkeit, intersubjektive Nachvollziehbarkeit und Kumulierbarkeit wissenschaftlicher Ergebnisse (Dafoe, 2014) gelegt werden kann, umfasst Open Science eine transparente Dokumentation des gesamten Prozesses empirischer Forschung, von der Datenerhebung bis zur Analyse und Veröffentlichung von Ergebnissen. Die Replikations- und Vertrauenskrisen, die verschiedenste wissenschaftliche Disziplinen in den letzten Jahrzehnten erfassten (Fraser, Parker, Nakagawa, Barnett, & Fidler, 2018; Ioannidis, 2005; Smith, 2019) und die auch die Sozialwissenschaften allgemein sowie die Politikwissenschaft im Besonderen miteinschließen (Burlig, 2018; Freese & Peterson, 2017; Monogan, 2013; Wuttke, 2019), machen deutlich, dass eine stärkere Ausrichtung der wissenschaftlichen Praxis auf Transparenz notwendig ist, um zu ermöglichen, dass Ergebnisse nachvollzogen, repliziert und revidiert werden können.

Auch GESIS - Leibniz-Institut für Sozialwissenschaften hat sich, neben vielen anderen wissenschaftlichen Akteuren und Initiativen, (z.B. Allianz der deutschen Wissenschaftsorganisationen, 2010; Deutsche Forschungsgemeinschaft, 2017) diesem Ziel bekannt und setzt sich für die Etablierung von Open Science Prozessen in den Sozialwissenschaften ein - u.a. durch die zur Verfügungstellung von Daten und deren Dokumentation für die wissenschaftlichen Community, dem Social Science Open Access Repository (SSOAR), in dem eine Vielzahl sozialwissenschaftlicher Dokumente als Volltexte kostenlos und frei zur Verfügung stehen, und SowiDataNet|Datorium, einer Plattform, auf der Wissenschaftler*innen ihren Code und ihre Daten mit der Community teilen können (für weitere Informationen siehe (GESIS - Leibniz-Institut für Sozialwissenschaften, 2020)).

Es ist unser Anliegen, die in der GESIS Open Science Strategie formulierten Ansprüche und Ziele mittel- bis langfristig in der German Longitudinal Election Study (GLES) umzusetzen und diese somit mehr auf die sich zunehmend in der Wissenschaft etablierenden Standards einer transparenten Wissenschaftspraxis auszurichten. In diesem Dokument sollen verschiedene Möglichkeiten skizziert werden, wie in der GLES die Ziele der GESIS übergreifenden Strategie umgesetzt und in diesem Rahmen erweitert werden können.

1.1 Open Science als Chance für die GLES

Die GLES als zentrale Infrastruktur für die kontinuierliche Erhebung und Bereitstellung qualitativ hochwertiger Daten für die nationale und internationale Wahlforschung leistet durch die schnelle Bereitstellung und freie Zugänglichkeit der Daten für die wissenschaftliche Community über den Datenbestandskatalog (DBK) bereits einen großen Beitrag zur praktischen Umsetzung von Open Science Grundsätzen. Jedoch könnte das Potential der GLES in Bezug auf die Umsetzung und Förderung von Open Science Praktiken noch stärker ausgeschöpft werden. Zum einen, indem weiter an der Erleichterung des Datenzugangs gearbeitet wird und zum anderen, indem Wissenschaftler*innen aktiv dabei unterstützt werden, ihre Forschung mit Daten der GLES transparent, d.h. nachvollziehbar, replizierbar und widerlegbar, zu machen. Darüber hinaus könnte durch eine entsprechende Publikationsstrategie die Umsetzung transparenter Forschung mit GLES-Daten incentiviert und so ein erheblicher Beitrag zur Etablierung von Open Science in der Wahlforschung geleistet werden.

Die Implementierung einer Open Science Strategie der GLES schärft zudem das Profil der Deutschen Wahlstudie und steigert ihren Bekanntheitsgrad. Sie folgt damit dem Beispiel der American National Election Study (ANES)² und profiliert sich somit auch im internationalen Kontext anderer Wahlstudien. Darüber hinaus zeigen verschiedene Entwicklungen der letzten Jahre³, dass es sich bei Forderungen nach mehr Replizierbarkeit, Reproduzierbarkeit und mehr Möglichkeiten zur Präregistrierung von empirischen Studien längst nicht mehr nur um eine Randerscheinung im wissenschaftlichen Arbeiten handelt. Vielmehr entwickelt sich die Umsetzung von Open Science Grundsätzen zu einer zentralen Forderung einer Bewegung, die sich anschickt, Forschungspraktiken nachhaltig zu verändern. Um diesen Anforderungen gerecht zu werden, wird auch ein transparenter und frühzeitiger Zugriff auf Informationen zur Erhebung von Daten und ihrer Veröffentlichung benötigt. Für die GLES bieten sich verschiedene Möglichkeiten, Beiträge zu einer transparenten und verlässlichen Wahlforschung zu leisten und die Sichtbarkeit der Studie in der Community zu erhöhen.

Im Folgenden wird vorgestellt, wie Arbeitsprozesse in der GLES strukturiert werden können, um diese Ziele zu erreichen. Die Umsetzung dieser Ziele ist als langfristiger Prozess angedacht, bei dem sich oftmals bereits existierenden Arbeitsprozesse noch stärker an einer offenen Wissenschaftspraxis orientieren sollen. Die Analyse von Verbesserungspotenzialen orientiert sich hierbei an den vier Kernbereiche von Open Science: Open Methodology (die transparente und reproduzierbare Dokumentation der Datenerhebung – und Analyse), Open Data (Zugänglichkeit der Forschungsdaten), Open Access (nachhaltige Verfügbarkeit von wissenschaftlichen Ergebnissen) und Open Source (Verwendung von nicht proprietärer Software).

² Für weitere Informationen siehe: <https://www.ercp2016.com/>

³ Beispielsweise erschien letztes Jahr ein Sonderheft der *Swiss Political Science Review* über „The Use and Usefulness of p-values in Political Science“ Bischof and Velden (2019); Troeger (2019).

2 Open Methodology

Open Methodology bezieht sich auf die Nachvollziehbarkeit und Reproduzierbarkeit der Methoden des gesamten Forschungsprozesses von der Planung, über die Datenerhebung bis hin zur Datenaufbereitung, Datenanalyse und Publikation der Ergebnisse. Für die GLES lassen sich zwei Bereiche unterscheiden, in denen Open-Methodology-Prinzipien relevant sind:

- (a) die Planung und Durchführung der Datenerhebung und Datenaufbereitung im Rahmen der GLES Studien und
- (b) die Ermöglichung und Förderung der Anwendung von Open Methodology Prinzipien durch Forschende, die GLES Daten nutzen.

Eine Abdeckung beider Bereiche lässt sich durch die Umstrukturierung der Dokumentation der Datenerhebung und Aufbereitungen in der GLES erreichen.

Darüber hinaus soll die Präregistrierung von Hypothesentests anhand von GLES Daten ermöglicht werden. Inwieweit Forschende ihren Forschungsprozess transparent gestalten können, hängt maßgeblich auch von Rahmenbedingungen ab, die Datenproduzierende setzen. Dies wird vor allem deutlich mit Blick auf die Präregistrierung wissenschaftlicher Studien. Diese meint, dass Studienvorhaben im Detail (Hypothesen, Analysepläne, etc.) öffentlich hinterlegt werden, bevor Zugriff auf die für die Analyse verwendeten Daten erfolgt (Nosek, Ebersole, DeHaven, & Mellor, 2018). Dieses Verfahren trägt insofern zur Etablierung einer transparenten wissenschaftlichen Praxis bei, als es eine klare Trennung zwischen der Formulierung und Überprüfung von Hypothesen erforderlich macht, da der Datenzugriff nachvollziehbar erst nach der Hypothesenformulierung erfolgt. Vor wenigen Jahren innerhalb der Disziplin noch unbekannt, hat das Instrument der Studienpräregistrierung innerhalb kurzer Zeit stark an Bedeutung gewonnen, sodass mittlerweile mehr als jede*r fünfte Politikwissenschaftler*in mindestens eine empirische Studie präregistriert hat (Christensen et al., 2019). Präregistrierungspläne sind vor allem in experimenteller Forschung verbreitet, finden aber auch mit Beobachtungs- und Sekundärdaten zunehmend Verwendung (Graham, Huber, Malhotra, & Mo, 2019).⁴ Voraussetzung für die Präregistrierung einer Studie ist dabei immer, dass sowohl die inhaltliche, als auch methodische Ausgestaltung einer Umfrage vor dem Zugriff auf die Daten bekannt sind.

Dass die Praxis der Präregistrierung sich bisher vor allem im Bereich der experimentellen Forschung etablieren konnte, liegt daran, dass die organisatorischen Voraussetzungen für eine (sinnvolle) Präregistrierung bei Umfragestudien sehr viel höher sind. Die GLES möchte einerseits demonstrieren, dass solche Bedingungen im Sinne einer offenen Wissenschaft geschaffen werden können. Andererseits wollen wir damit einem sehr großen Kreis weiterer Forschenden die Möglichkeit geben, ihre

⁴ Präregistrierung hilft analysis bias' zu minimieren (Jacobs, 2018), indem es die nachträgliche Anpassung der Analysemethode an die Analysebefunde unterbindet oder zumindest transparent macht. Um 'publication bias' – also der Beschränkung von Veröffentlichungen in peer reviewed Journals auf signifikante Befunde – zu minimieren findet ein neues Artikelformat in den Sozialwissenschaften seit einigen Jahre stärkere Verbreitung: Registered Reports setzen auf das Instrument der Präregistrierung auf und verbinden es mit ergebnisblinden Review-Verfahren. Indem Studien vor der Datenerhebung begutachtet und zur Veröffentlichung angenommen oder abgelehnt werden, rückt die Relevanz der Forschungsfrage und die theoretischen und methodischen Ansätze zur Bearbeitung dieser Frage als ausschlaggebendes Kriterium in den Mittelpunkt (für weitere Informationen hierzu siehe (Jacobs, 2018) oder auch <https://cos.io/rr/>). Mittlerweile haben 244 Fachzeitschriften Registered Reports eingeführt, einschließlich einer wachsenden Zahl an sozial- und politikwissenschaftlichen Journalen (Research and Politics, Politics and Life Science, Nature Human Behavior, Journal of Experimental Political Science).

individuellen Forschungsprozesse stärker an Open Science Prinzipien auszurichten. Wir sind davon überzeugt, dass die Schaffung von Möglichkeiten zur Präregistrierung das wissenschaftliche Gewicht der Ergebnisse, die mithilfe der qualitativ hochwertigen Daten der GLES erzielt werden können, noch zusätzlich zu steigern vermag.

2.1 Status Quo

Im Bereich der Open Methodology ermöglicht die GLES durch die umfassende Dokumentation des Frageprogramms sowie der Datenerhebung und Datenaufbereitung die Überprüfbarkeit und Reproduzierbarkeit ihrer Methoden. Zudem ist die gesamte Datenaufbereitung syntaxbasiert und reproduzierbar.

2.2 Verbesserungspotenzial

(a) Dokumentation von Befragungsinstrumenten und Analysen

Während die Methodiken der Erhebung und Aufbereitung der GLES Daten sehr transparent dokumentiert sind, könnte die Entwicklung der Befragungsinhalte besser dokumentiert sein. Dazu müsste an geeigneter Stelle berichtet werden, ob Skalen oder Einzelfragen aus anderen Quellen übernommen, adaptiert oder selbst entwickelt wurden. Zudem werden bei der Dokumentation von Befragungsinhalten gelegentlich Label verwendet, die nur schwer Rückschlüsse auf die sich dahinter verbergenden Konstrukte erlauben, so dass es für Nutzerinnen und Nutzer schwierig sein kann, zu erkennen ob, wo und wann spezifische Inhalte und Konstrukte in der GLES erhoben wurden.

In der GLES werden in regelmäßigen Abständen Datenqualitätsberichte erstellt, die detaillierte Informationen über die Zusammensetzung der realisierten Stichprobe sowie zur Ausschöpfung beinhalten und auch weitere Indikatoren zur Bewertung der Qualität der Daten heranziehen. Die Codes bzw. Syntaxen der Analysen und Grafiken dieser Berichte werden derzeit in der Regel nicht publiziert, was aber die Transparenz der Methodik weiter erhöhen und eine weitere Dienstleistung gegenüber der wissenschaftlichen Community darstellen würde.

Generell setzen Mitarbeitende des GLES Teams Open Methodology Prinzipien mehr und mehr bei eigenen Publikationen und Forschungsprojekten um. Um die GLES in diesem Kontext als Vorreiter zu etablieren und die Außenwirkung zu steigern, kann dieses Bemühen noch weiter unterstützt werden, indem die Einhaltung von Open Methodology Prinzipien explizit angestrebt wird.

(b) Ermöglichung und Förderung der Anwendung von Open Methodology Prinzipien durch Forschende, die GLES Daten nutzen.

Durch die zeitgleiche Veröffentlichung von Daten und Fragebogendokumentationen in der GLES sind zwar viele Schritte des Forschungsprozesses zum Zeitpunkt der Datenveröffentlichung nachvollziehbar und transparent. Allerdings macht dieser zeitliche Ablauf eine Präregistrierung oder das Verfassen eines Registered Reports mit GLES Daten im deutschsprachigen als auch im englischsprachigen Kontext für Datennutzende nicht möglich.

2.3 Maßnahmen

(a) Dokumentation von Befragungsinstrumenten und Analysen

2.3.1 Publikationen zu spezifischen Inhalten des GLES Fragenprogramms

Wir werden erproben, die Messung und Adressierung verschiedener Inhalte, Themen und Konstrukte in der GLES in themenspezifischen Publikationen zu dokumentieren, übersichtlich darzustellen und zu erläutern. Hierbei sollen auch die Herkunft und die theoretischen Hintergründe einzelner Instrumente erläutert, sowie die Qualität einzelner Maße berichtet werden. Um eine höhere Reichweite zu erzielen und neue Nutzungsgruppen für die GLES zu erschließen, sollen diese Publikationen in englischer Sprache verfasst werden.

2.3.2 Publikation der Codes von Analysen

Bei Publikationen von Dokumentationen der Studien der GLES (z.B. Qualitätsberichten) wird in Zukunft auch der Code der Analysen frei zugänglich publiziert. Die Publikationen enthalten entsprechende Verweise auf den Code.

Mitarbeitende der GLES stellen in Zukunft bei allen eigenen Publikationen, die auf Daten der GLES basieren, den Code der Analysen nach Möglichkeit frei zugänglich zur Verfügung.

(b) Ermöglichung und Förderung der Anwendung von Open Methodology Prinzipien durch Forschende, die die GLES Daten nutzen

2.3.3 Veröffentlichung des Frageprogramms der GLES mit dem Feldstart

Um Forschenden die Möglichkeit zu geben, Forschungsfragen mit Daten der GLES zu präregistrieren oder Registered Reports anzufertigen und frühzeitig ihre Studien zu planen, werden die Fragebögen der GLES in der Regel zum Feldstart auf der GLES Homepage veröffentlicht. Neben dem Frageprogramm enthalten diese auch methodische Vorab-Informationen zur Erhebung. Dieser Schritt wurde bereits durch die Koordinationsgruppe beschlossen und wird im Jahr 2020, seit der 13. Welle des GLES Panels bereits umgesetzt. Mit der Fragebogenveröffentlichung wird dabei auch bereits transparent kommuniziert, wann die Datenveröffentlichung geplant ist. Das Zeitfenster, in dem die Datenveröffentlichung stattfinden soll, wird dabei zur besseren Planbarkeit möglichst klein gehalten (ca. 4 Wochen).

2.3.4 Veröffentlichung von englischen Übersetzungen des Frageprogramms der GLES mit dem Feldstart

Um die Zugänglichkeit unserer Daten nicht durch sprachliche Barrieren zu erschweren, sowie die Möglichkeit der Präregistrierung mit GLES Daten auf die internationale Ebene zu transferieren, soll die englische Übersetzung des Frageprogramms bereits frühzeitig im Workflow platziert werden, sodass eine Veröffentlichung des englischen Fragebogens ebenfalls wenn möglich zum Feldstart erfolgen kann. Hierfür wird der Prozess der Übersetzung optimiert, um diesen einerseits zeiteffizient zu gestalten, andererseits aber auch ein gewisses Maß an Flexibilität zu gewährleisten. Auch dies wird bereits ab 2020 versucht umzusetzen, die 13. Welle des GLES Panels fungierte hier als Pilotprojekt, um die Prozesse noch weiter anzupassen. Da für das Frageprogramm der GLES 2009-2017 bereits englische Übersetzungen vorliegen, kann hier größtenteils auf bereits geleistete Vorarbeit zurückgegriffen werden. Im Hinblick auf die Veröffentlichung englischer Frageprogramme soll jedoch auch weiterhin und für neue Items eine hohe Qualität der Übersetzung gewährleistet werden. Deshalb unterlaufen neue Items vor Veröffentlichung ein Lektorat, um sicherzustellen, dass die ursprüngliche Bedeutung von Begrifflichkeiten auch im Englischen erhalten bleibt. Um eine Vorabveröffentlichung einer englischen Übersetzung des Fragebogens zu Feldstart zu ermöglichen, ist es deshalb notwendig, das Kernprogramm der jeweiligen Erhebungen möglichst frühzeitig zu platzieren.

2.3.5 GLES Präregistrierungs-Challenge

Wir schlagen vor, die Präregistrierung von Studien aktiv zu bewerben und zu incentivieren. Ein Beispiel hierfür ist die Durchführung einer Präregistrierungs-Challenge. Wie oben beschrieben, bedeutet Präregistrierung zunächst einmal, dass Theorien, Hypothesen und empirische Modelle festgelegt werden, bevor der Zugriff auf die Daten erfolgt. Voraussetzung für die Präregistrierung ist somit die Vorab-Veröffentlichung von deutschen und englischsprachigen Fragebögen für die internationale Wissenschafts-Community. Um eine solche Präregistrierung zu incentivieren und auch die GLES als Vorreiter der Open Science Initiative zu etablieren, schlagen wir die Planung eines Special Issues zu einer Bundestagswahl, herausgegeben in Kooperation der KG GLES und des GLES Teams bei GESIS, vor. Verbunden mit einer Präregistrierungs-Challenge würden ausschließlich präregistrierte Studien eingereicht werden können und anschließend einen möglichst transparenten Review-Prozess durchlaufen.

3 Open Data

Die Erhebung und Bereitstellung von qualitativ hochwertigen Daten für die nationale und internationale Wahlforschung sind die Kernaufgaben der GLES. Die Bereitstellung der GLES Daten erfolgt über das GESIS Datenarchiv für Sozialwissenschaften, das sich der Idee offener Daten seit seiner Gründung verschrieben hat. Die FAIR Prinzipien (Eder & Jedinger, 2019; Wilkinson et al., 2016) geben darüberhinaus Datenproduzenten einen Kriterienkatalog für eine bessere maschinelle Verarbeitbarkeit von Daten und Metadaten an die Hand. Da nicht nur immer mehr, sondern auch in zunehmender Geschwindigkeit Daten veröffentlicht werden, sollten Daten nach diesen Kriterien *findable*, *accessible*, *interoperable* und *reusable* sein. Zwar sind FAIR Data nicht immer auch frei zugänglich im Sinne der Open Data, aber die Umsetzung der FAIR Prinzipien trägt erheblich zur (Nach-)Nutzung der Daten bei.

3.1 Status Quo

Im GESIS Datenarchiv sind neben den Daten und Metadaten selbst auch detaillierte Beschreibungen der Studien, die Fragebögen und weitere begleitende Dokumente verfügbar. Außerdem sind Errata und die Versionsverläufe der Datensätze und Dokumente dokumentiert. Ein Direktdownload der Dokumentationen der GLES Daten ist ohne Registrierung möglich. Der Großteil der Datensätze der GLES kann zudem nach einer einmaligen Registrierung von Forschende kostenfrei heruntergeladen und genutzt werden. Das GLES Team hat sich dabei bereits frühzeitig dazu entschieden, den Zugang zu einzelnen Variablen oder Datensätze nur in Ausnahmefällen, die sich insbesondere durch die Wahrung des Datenschutzes begründen, zu beschränken. Um auch Zugang zu datenschutzsensiblen Datensätzen oder einzelnen Variablen zu ermöglichen, können diese Informationen abhängig von der Sensibilität der Daten nach Abschluss eines Datennutzungsvertrages bereitgestellt (off-site Nutzung) oder vor Ort im Secure Data Center analysiert werden (on-site Nutzung).

3.2 Verbesserungspotenzial

In einer vergleichenden Untersuchung nationaler Wahlstudie haben Eder & Jedinger (2019) 18 Wahlstudien mit Blick auf die Umsetzung der FAIR Prinzipien evaluiert (Eder & Jedinger, 2019). Dabei schneidet die GLES vor allem in der Qualität und der Dokumentation der Daten gut ab und es wird hohe Interoperabilität gewährleistet. Verbesserung könnten aber vor allem in den Bereichen der Accessibility und Reusability erreicht werden.

Zum einen liegen für einzelne Datensätze, die nicht zum Kernprogramm gehören (z.B. Pretests), und deren Dokumentationen keine englische Übersetzungen vor oder diese werden erst mit Verzögerung veröffentlicht, da sie einen aufwendigen Lektoratsprozess durchlaufen. Daneben entspricht auch die Zugänglichkeit der Daten durch die standardmäßige Veröffentlichung in proprietären Datenformaten (SPSS und Stata) nicht gänzlich den Open Access Anforderungen (siehe auch Open Source).

3.3 Maßnahmen

3.3.1 Veröffentlichung von englischen Übersetzungen der Datensätze

Durch die Korrespondenz von Fragebogen, Codierungen und Variablennamen im Datensatz können auch englischsprachige Forschende die Daten direkt zum Veröffentlichungszeitpunkt nutzen. Es wird langfristig angestrebt, englische Übersetzungen der Daten und weiterer Dokumente frühzeitig zu veröffentlichen, um englischsprachigen Forschenden die Nutzung der GLES Daten zu erleichtern und somit zu erreichen, dass die GLES vermehrt auch in der internationalen Community genutzt wird. Durch die Harmonisierung des Kernfragebogens ist dabei bereits ein erster Schritt getan, da die Übereinstimmung zwischen den Studien der GLES zunimmt und somit auch auf bereits existierende Übersetzungen zurückgegriffen werden kann. Allerdings wird der Aufwand für die Übersetzung von Studiendokumentationen und Datensätzen für das Bundestagswahljahr 2021 relativ groß ausfallen, so dass hier noch mit einer gewissen Verzögerung zu rechnen ist.

4 Open Access

Unter Open Access wird üblicherweise die kostenfreie Zugänglichkeit von Forschungspublikationen im Internet verstanden. Nach unserem Verständnis sollte Open Access sich jedoch nicht nur auf rechtliche und finanzielle Hürden der Zugänglichkeit von Forschungsergebnissen beschränken, sondern auch sprachliche Hürden und Barrieren im Blick haben, die oftmals Menschen mit Behinderung den Zugang zu Forschungsergebnissen erschweren oder unmöglich machen.

4.1 Status Quo

Im Punkt Open Access herrscht in der GLES noch Verbesserungspotential. Publikationen aus der GLES sind häufig nicht Open Access verfügbar (siehe PopScience Buch oder viele Aufsätze in Fachzeitschriften).

Publikationen im Rahmen der GLES werden bisher nicht systematisch hinsichtlich Barrierefreiheit geprüft.

4.2 Verbesserungspotenzial

Es wäre wünschenswert, wenn zukünftig möglichst viele Publikationen, die auf Daten der GLES beruhen, frei zugänglich sind. Dazu werden insbesondere Mitarbeitende der GLES bei GESIS aufgefordert, bei Publikationen die Möglichkeit von *Open Access* zu prüfen und nach Möglichkeit zu nutzen.

4.3 Maßnahmen

4.3.1 Open Access Publikationsstrategie

Mitarbeitende der GLES bei GESIS werden eigene Forschungsergebnisse, die auf Daten der GLES basieren, nach Möglichkeit frei im Netz verfügbar machen, indem sie diese in genuinen Open Access Medien publizieren (goldener Weg), bei Medien mit hybriden Open Access Modellen den spezifischen Artikel frei verfügbar machen, oder im Rahmen des Zweitveröffentlichungsrechts eine Version des Artikels in einem Repositorium archivieren (grüner Weg). GESIS stellt im Rahmen seiner Open Access Policy⁵ bis zu einem bestimmten Umfang Mittel zur Verfügung, die von Beschäftigten zur Begleichung von Publikationsgebühren beantragt werden können. Darüber hinaus betreibt GESIS das Social Science Open Access Repository (SSOAR), welches der langfristigen Archivierung von Publikationen dient.

4.3.2 Hervorhebung von Open Access Publikationen mit GLES Daten

Bei der Anzeige der Ergebnisse von Suchen in der GLES Bibliografie werden Open Access Publikationen in Zukunft besonders hervorgehoben.

4.3.3 Gewährleistung der Barrierefreiheit von GLES Publikationen

GLES Publikationen werden zukünftig im Hinblick auf Barrierefreiheit weiter optimiert.

⁵ <https://www.gesis.org/angebot/publikationen/open-access-policy/>

5 Open Source

Open Source bezieht sich auf die freie Verfügbarkeit des Codes eingesetzter Software unter Umständen auch weiterentwickelt und modifiziert werden kann. Open Source Software ist in vielen Fällen kostenlos verfügbar und der Quellcode kann eingesehen werden.

5.1 Status Quo

Die Daten der GLES Studien werden in den proprietären Formaten der Statistiksoftware Stata und SPSS veröffentlicht.

5.2 Verbesserungspotenzial

Obwohl Stata und SPSS an vielen wissenschaftlichen Einrichtungen Verwendung finden, nimmt die Nutzergruppe von frei zugänglichen Programmen wie beispielsweise R und Python kontinuierlich zu. Bei der Umwandlung von Datensätzen aus proprietären Dateiformaten kann es zu Einlese- und Darstellungsfehlern kommen, die sich im schlimmsten Fall auf Analyseergebnisse auswirken könnten.

5.3 Maßnahmen

5.3.1 Bereitstellung der Daten im CSV-Dateiformat

Um dafür zu sorgen, dass möglichst alle interessierten Forschende mit den GLES Daten arbeiten können, werden wir dafür sorgen, dass die Nutzung von GLES Daten und deren Dokumentation nicht an den Besitz der Lizenzen bestimmter Softwareprogramme gebunden ist. Somit werden Datensätze neben SPSS- und Stata-Formaten auch in allgemein nutzbaren und frei zugänglichen Datenformaten, wie beispielsweise CSV, veröffentlicht. Neben der freien Zugänglichkeit hat das Dateiformat CSV zudem den Vorteil, dass eine langfristige Archivierung der Datensätze, welche auch eine langfristige Replizierbarkeit von Studien, die auf unsere Datensätze zurückgreifen, impliziert, gewährleistet werden.

References

- Allianz der deutschen Wissenschaftsorganisationen (2010). Grundsätze zum Umgang mit Forschungsdaten. Retrieved from <https://www.leopoldina.org/fileadmin/redaktion/Publikationen/Allianz/100624-allianz-forschungsdaten.pdf>
- Bischof, D., & Velden, M. (2019). The Use and Usefulness of p-Values in Political Science: Introduction. *Swiss Political Science Review*, 25(3), 276–280. <https://doi.org/10.1111/spsr.12376>
- Burlig, F. (2018). Improving transparency in observational social science research: A pre-analysis plan approach. *Economics Letters*, 168, 56–60. <https://doi.org/10.1016/j.econlet.2018.03.036>
- Christensen, G., Wang, Z., Paluck, E. L., Swanson, N., Birke, D. J., Miguel, E., & Littman, R. (2019). *Open Science Practices are on the Rise: The State of Social Science (3S) Survey*. <https://doi.org/10.31222/osf.io/5rksu>
- Dafoe, A. (2014). Science Deserves Better: The Imperative to Share Complete Replication Files. *PS: Political Science & Politics*, 47(01), 60–66. <https://doi.org/10.1017/S104909651300173X>
- Deutsche Forschungsgemeinschaft (2017). Replizierbarkeit von Forschungsergebnisse: Eine Stellungnahme der Deutschen Forschungsgemeinschaft. Retrieved from https://www.dfg.de/download/pdf/dfg_im_profil/reden_stellungnahmen/2017/170425_stellungnahme_replizierbarkeit_forschungsergebnisse_de.pdf
- Eder, C., & Jedinger, A. (2019). FAIR national election studies: How well are we doing? *European Political Science*, 18(4), 651–668. <https://doi.org/10.1057/s41304-018-0194-3>
- European Union (2019). Open Science. Retrieved from https://ec.europa.eu/info/sites/info/files/research_and_innovation/knowledge_publications_tools_and_data/documents/ec_rtd_fact-sheet-open-science_2019.pdf
- Fraser, H., Parker, T., Nakagawa, S., Barnett, A., & Fidler, F. (2018). Questionable research practices in ecology and evolution. *PLoS One*, 13(7), e0200303. <https://doi.org/10.1371/journal.pone.0200303>
- Freese, J., & Peterson, D. (2017). Replication in Social Science. *Annual Review of Sociology*, 43(1), 147–165. <https://doi.org/10.1146/annurev-soc-060116-053450>
- GESIS - Leibniz-Institut für Sozialwissenschaften (2020). GESIS OPEN Science Strategy.
- Graham, M. H., Huber, G. A., Malhotra, N., & Mo, C. H. (2019). *Observational Open Science: An Application to the Literature on Irrelevant Events and Voting Behavior*. <https://doi.org/10.31222/osf.io/udfs9>
- Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*, 2(8), e124. <https://doi.org/10.1371/journal.pmed.0020124>
- Jacobs, A. M. (2018). *Of Bias and Blind Selection: Pre-registration and Results-Free Review in Observational and Qualitative Research*.
- Monogan, J. E. (2013). A Case for Registering Studies of Political Outcomes: An Application in the 2010 House Elections. *Political Analysis*, 21(1), 21–37. <https://doi.org/10.1093/pan/mps022>
- Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences of the United States of America*, 115(11), 2600–2606. <https://doi.org/10.1073/pnas.1708274114>
- Smith, N. (2019, September 17). Why economics is having a replication crisis: Recreating research by gathering data from the real world and analyzing it statistically often fails to produce the same result. *Bloomberg Opinion*. Retrieved from <https://www.bloomberg.com/opinion/articles/2018-09-17/economics-gets-it-wrong-because-research-is-hard-to-replicate>

- Troeger, V. E. (2019). To P or not to P? The Usefulness of P-values in Quantitative Political Science Research. *Swiss Political Science Review*, 25(3), 281–287. <https://doi.org/10.1111/spsr.12377>
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J. J., Appleton, G., Axton, M., Baak, A., . . . Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3, 160018. <https://doi.org/10.1038/sdata.2016.18>
- Wuttke, A. (2019). Why Too Many Political Science Findings Cannot Be Trusted and What We Can Do About It: A Review of Meta-Scientific Research and a Call for Academic Reform. *Politische Vierteljahresschrift*, 60(1), 1–19. <https://doi.org/10.1007/s11615-018-0131-7>