**Forschungsdatenmanagementplan (FDMP) – Skizzenphase**

|  |  |
| --- | --- |
| 1. **Administrative Angaben** | |
| Projekttitel | *Titel des geplanten Forschungsprojekts mit Akronym* |
| Skizzeneinreicher | *Name des Skizzeneinreichers* |
| Verbundpartner | *Namen aller Verbundpartner* |
| Förderprogramm | *Bsp.: Programm zur Innovationsförderung des Bundesministeriums für Ernährung und Landwirtschaft* |
| Bekanntmachung | *Bsp.: Bekanntmachung über die Förderung von Innovationen zur Züchtung von klimaangepassten Sorten und Kulturpflanzen im Rahmen des Programms zur Innovationsförderung* |
| Datum | *Datum der Erstellung des FDMP* |
| **1.a Policy & Leitlienie**  Sämtliche Datenverarbeitung am ZALF unterliegt der „Leitlinie zum Umgang mit Forschungsdaten in der Leibniz-Gemeinschaft“ (<https://www.leibniz-gemeinschaft.de/fileadmin/user_upload/Bilder_und_Downloads/Forschung/Open_Science/Leitlinie_Forschungsdaten_2018.pdf>)  Die ZALF Data Policy (<https://doi.org/10.5281/zenodo.2567723>) definiert relevante Begriffe, Rechte und Lizenzen, den Umgang mit Forschungsdaten durch einen ausführlichen Datenmanagementplan und Verantwortlichkeiten. Der Datenschutzbeauftragte des ZALF ist in diesen Prozess eingebunden ([datenschutz@zalf.de](mailto:datenschutz@zalf.de)). | |
| 1. **Beschreibung der Datenerhebung und -verarbeitung** | |
| Die folgenden Daten (Parameter, Spalten mit Werten) werden erhoben, verarbeitet und veröffentlicht:   * Experimentaldaten: Messwerte, … * Simulationen: vollständige Inputdateien, Ergebnisdateien, ... * Interviewdaten: Fragestellung, Zielgruppe, … * Sekundärdaten: Inhalte, Quellen, zugehörige Verarbeitungsskripte   Die folgenden Daten Dritter werden im Projekt nachgenutzt:  …  …  Eine Weiterverarbeitung der Daten erfolgt im Wesentlichen mit Hilfe von \_\_\_hier bitte die verwendeten aktuellen Methoden der Datenverarbeitung aufführen\_\_\_\_\_\_.  Verwendete Skripte sind Bestandteil der Forschungsarbeit und werden entsprechend der guten Wissenschaftlichen Praxis gemeinsam mit den Daten veröffentlicht.  Die oben genannten Daten werden in den folgenden Formaten gespeichert und weiterverarbeitet: \_\_\_\_\_\_\_bitte die vorgesehenen Formate benennen\_\_\_\_\_\_\_.  Diese Formate sind mit dem BonaRes Repositorium kompatibel und einer Veröffentlichung der Daten und damit der Nachnutzung steht nichts im Wege. Wir gehen davon aus, dass sich das Volumen der im Projekt erhobenen Daten auf \_\_\_\_\_XXTB\_\_\_\_ und das Volumen der veröffentlichten Daten auf \_\_\_\_\_\_\_XXTB\_\_\_\_\_\_ beschränkt. | |
| 1. **Dokumentation der Daten** | |
| Alle Daten werden nach Erhebung strukturiert gespeichert und mit standardisierten Metadaten nach dem BonaRes Metadatenschema (<https://doi.org/10.20387/bonares-5pgg-8yrp>) beschrieben. Die Metadaten beinhalten den DataCite (<https://schema.datacite.org/meta/kernel-4.1>) und INSPIRE (<https://inspire.ec.europa.eu/metadata/6541>) Standard und sind durch disziplinspezifische Metadatenelemente erweitert, um die wissenschaftliche Nachnutzung zu erleichtern. In den Metadaten werden die kontollierten Vokabulare [AGROVOC](https://www.fao.org/agrovoc/) (FAO) und [GEMET](https://www.eionet.europa.eu/gemet/en/themes/) (EEA) verwendet, um eine internationale Interoperabilität sicherzustellen. Die gesamte Historie der Datenerzeugung, Verarbeitung, Bereitstellung und eventueller nachträglicher Veränderungen (Updates) werden als „Lineage“ in den Metadaten festgehalten.  Die Datenqualität wird im Laufe der Datenerhebung und Verarbeitung durch die folgenden Methoden überwacht und sichergestellt: \_\_\_\_\_bitte ergänzen\_\_\_\_\_\_  Im Zuge der Veröffentlichung im BonaRes Repositorium werden formale Aspekte der Datenqualität (Struktur, Format) und die Metadaten kontrolliert und ggf. nachgebessert. Die im BonaRes Repositorium vorliegenden Expertise auf dem Gebiet der Kuration agrarwissenschaftlicher Daten wird genutzt, um die Daten inhaltlich auf Plausibilität zu prüfen. Daten werden, im Sinne einer einfachen Nachnutzung, in verbreiteten, nicht proprietären Formaten bereitgestellt (CSV, TXT, XLSX, GDB); spezielle Tools oder Software sind nicht nötig (\_\_\_\_\_prüfen, ob dies gewährleitet ist\_\_\_\_\_\_). | |
| 1. **Datenselektion und Datenarchivierung** | |
| Die Daten, die aus diesem Projekt hervorgehen, werden am ZALF gespeichert und unterliegen dem deutschen Recht. Die Datenspeicherinfrastruktur des ZALF basiert auf einem mehrstufigen Ansatz. Aktive Daten werden auf  hochleistungsfähigen Flash-Disks gespeichert (ZALF Storage), die einen sofortigen und schnellen Zugriff ermöglichen. Primär Daten werden mehrfach täglich durch s.g. Snaphots gesichert, die eine schnelle Wiederherstellung ermöglichen. Die Primärdaten werden durch Spiegelung in einem räumlich getrennten Rechenzentrum (auf dem Gelände des ZALF) für Datenrettungszwecke repliziert. Dies ermöglicht im Störungsfall, den Weiterbetrieb über das Spiegelrechenzentrum. Der maximale Datenverlust liegt bei bis zu < 5 min (Recovery Point Objektive: RPO); die Zeit, die für die Wiederherstellung notwendig ist, liegt bei < 2h während der Arbeitszeiten (Recovery Time Objektive: RTO). Bei Bedarf kann eine Verschlüsselung der Daten weitere Datensicherheit bringen (\_\_\_\_\_prüfen, ob dies notwendig ist und ggf, vorher mit der ZALF IT abstimmen. Alternative wählen und einfügen: a) Die vom ZALF angebotene Verschlüsselung der Daten für weitere Sicherheit ist für das Projekt nicht notwendig. Ggf. Begründung einfügen; b) Im Projekt werden sensible Daten durch eine Verschlüsselung zusätzlich gesichert. \_\_\_\_\_\_).  Zudem werden tägliche inkrementelle und wöchentlich vollständige Backups auf Magnetspeicher (Festplatte) gesichert, mit einer Aufbewahrungszeit von vier Wochen. Zusätzlich wird einmal monatlich ein Fullbackup auf Tape (Magnetband) gesichert.  Die Medien werden offline an einem separaten Ort gelagert und sind somit dem direkten Zugriff entzogen. Für diese Sicherung beträgt die RPO 12 Monate.  Der Datenzugriff auf alle am ZALF gespeicherten Daten ist durch ein umfangreiches und feingranulares Nutzermanagement abgesichert. | |
| 1. **Verfügbarmachung der Projektdaten** | |
| Rechtliche Besonderheiten im Umgang und der Veröffentlichung von Daten liegen nicht vor \_\_\_\_\_\_bitte prüfen\_\_\_\_\_\_\_  Starke Restriktionen für die Nachnutzung der Daten liegen nicht vor \_\_\_\_\_\_bitte prüfen\_\_\_\_\_\_\_  Alle Daten werden mit umfangreichen Metadaten beschrieben. Ein Teil dieser Metadaten regelt die Bedingungen der Nachnutzung (Zugangs- und Nutzungsrechte). Der ZALF Ansatz ist es, Daten möglichst frei bereitzustellen. Daher wird standardmäßig die CC-BY Lizenz der Creative Commons vergeben. Es besteht die Möglichkeit für ZALF Autoren ein Embargo zu vergeben und den Zugang zu den Daten damit zeitlich zu befristeten. Hiermit wird den Ansprüchen auf das Erstverwertungsrecht Rechnung getragen.  Wir Die unter (2.) beschriebenen Daten können zu folgenden Zwecken nachgenutzt werden (Beispiele): \_\_\_\_\_\_\_\_\_\_bitte ausführlich weitere Nutzungsmöglichkeiten angeben\_\_\_\_\_\_\_\_\_  verfolgen das Ziel erhobene Daten möglichst umfangreich zur Nachnutzung bereitzustellen. Die Veröffentlichung geschieht innerhalb von 24 Monaten nach der Erhebung und Qualitätssicherung mindestens aber vor Beendigung des Projektes. Weitere Spezifikationen sind in der Projekt Data Policy festgelegt. \_\_\_\_\_prüfen und anpassen\_\_\_\_\_ Das wichtigste Kriterium zur Veröffentlichung ist es, ausschließlich qualitätsgeprüfte Daten zu veröffentlichen (siehe (2.)). Wenn es sich um sensible Daten handelt, werden diese vor der Veröffentlichung ausreichend anonymisiert. Ungeprüfte Rohdaten werden nicht veröffentlicht. Alle Daten, die im BonaRes Repositorium veröffentlicht sind, werden permanent und für lange Zeit verfügbar sein. Die dem BonaRes Repositorium zu Grunde liegende Datenbank ist vollständig in die ZALF Speicherinfrastruktur integriert (siehe (4.)) | |

**Ausfüllhinweise**

Generelle ZALF FDM Hinweise

Der vorliegende DMP enthält bereits grundlegende Informationen zum ZALF Datenmanagement, dient als Vorlage und ist an die jeweiligen Anforderungen, die sich aus dem Projekt ergeben anzupassen.

Insbesondere die gelb markierten Hinweise müssen ergänzt und angepasst werden. Dieser DMP wird bei Förderzusage der ZALF AG FDM zur Verfügung gestellt.

zu: 2. Beschreibung der Datenerhebung und -verarbeitung

Beschreiben Sie bitte, welche Daten in dem geplanten Forschungsprojekt nachgenutzt oder neu erhoben werden sollen, wie die Daten erhoben, strukturiert, prozessiert und ausgewertet werden sollen, welche Methoden zur Qualitätssicherung angewendet werden sollen, mit welchem Datenvolumen gerechnet wird und ob die erhobenen Daten grundsätzlich nachnutzbar sind.

*Bsp.:*

*Nach aktuellen Recherchen in den fachspezifischen Datenarchiven (Repositorien) sind keine Daten verfügbar, die für die Fragestellungen im Projekt genutzt werden können. Aus diesem Grund sollen folgende Datenerhebungen erfolgen: …*

*Im ersten Versuchsjahr erfolgt im Entwicklungsstadium XY (BBCH XY) eine quantitative Bestimmung des Larvenbefalls an 20 Ähren /Parzelle/Linie. Die Daten werden mit den zugehörigen Metadaten als Excel-Datei abgelegt. Die Datenanalyse erfolgt mit Hilfe der Statistik-Software XY. Zur Qualitätssicherung der Daten werden die Bonituren im zweiten Versuchsjahr (identische Standorte und Wiederholungen) wiederholt. Das Volumen der in dem Projekt erhobenen Daten wird insgesamt auf …GB geschätzt.*

*Der Datensatz zur Untersuchung von XY kann für andere Forschungseirichtungen von Relevanz sein. Die Daten und Metadaten sollen daher für die Nachnutzung in einem fachspezifischen Repositorium bereitgestellt werden.*

zu: 3. Dokumentation der Daten

Erläutern Sie bitte, wie die Daten dokumentiert werden, damit sie nachnutzbar sind. Geben Sie, falls vorhanden, Standards an, die für die Datenspeicherung, -benennung, -versionierung und die Zusammenstellung der Metadaten genutzt werden.

*Bsp.1: Die Datenspeicherung, -benennung, -versionierung erfolgt nach den Standards der Universität XY. Die Datenspeicherung erfolgt über das Rechenzentrum der Universität XY. Die Metadaten werden nach den Vorgaben des Repositoriums XY zusammengestellt und sind damit maschinenlesbar und interoperabel.*

*Bsp.2: Die Standards zur Datenspeicherung, -benennung, -versionierung und die Zusammenstellung der Metadaten werden vor Projektbeginn mit den Verbundpartnern festgelegt und im Falle einer Antragseinreichung konkretisiert.*

zu: 4. Datenselektion und Datenarchivierung

Erläutern Sie bitte, welche Projektdaten für welchen Zeitraum archiviert werden sollen. Welche Daten sollen ggf. nicht archiviert werden und warum? Hinweis: Vornehmlich sind die Daten zu archivieren, die einer Publikation zugrunde liegen oder mit denen die wesentlichen Ergebnisse eines Projektes nachgewiesen oder reproduziert werden können.

Hinweis: Den Regeln der Guten Wissenschaftlichen Praxis folgend, sollten Forschungsdaten in der eigenen Einrichtung oder in einer fachlich einschlägigen, überregionalen Infrastruktur für mindestens zehn Jahre archiviert werden.

Machen Sie Angaben zum geschätzten Datenvolumen der zu archivierenden Daten und den damit verbundenen Ausgaben/Kosten während der Projektlaufzeit und danach.

*Bsp.: Die Projektergebnisse sollen in mindestens zwei OpenAccess Publikationen veröffentlicht werden. Alle hierfür relevanten Projektdaten, einschließlich der Metadaten, sollen in einem fachspezifischen Repositorium (z.B. Repositorium XY) für einen Zeitraum von zehn Jahren archiviert und bereitgestellt werden.*

*Während der Projektlaufzeit wird mit einem zu archivierenden Datenvolumen von …GB gerechnet. Die Ausgaben/Kosten für die Datenarchivierung belaufen sich während der Projektlaufzeit auf voraussichtlich…€. Nach Projektende wird das zu archivierende Datenvolumen auf …GB geschätzt. Die Ausgaben/Kosten hierfür werden von der Forschungseinrichtung getragen.*

zu: 5. Verfügbarmachung der Projektdaten

Erläutern Sie bitte welche Daten zu welchem Zeitpunkt veröffentlicht und zur Nachnutzung bereitgestellt werden sollen. In welchem Repositorium sollen die Daten archiviert werden und welche Zugangsmöglichkeiten werden eingeräumt (frei, eingeschränkt, kein Zugang, lizensiert)? Erläutern Sie ggf. welche Daten nicht veröffentlicht bzw. nicht zur Nachnutzung bereitgestellt werden. Es können bspw. rechtliche (z.B. patentrechtliche, urheberrechtliche), wettbewerbliche oder ethischen Aspekten vorliegen, aufgrund derer eine Bereitstellung der Daten nicht möglich ist.

*Bsp.1: Die Projektdaten werden mit der Veröffentlichung der Projektergebnisse, spätestens jedoch mit dem Schlussbericht sechs Monate nach Projektende, in dem institutionellen Repositorium der Universität XY frei zugänglich und mit einem Identifikator (DOI) versehen, zur Verfügung gestellt.*

*Bsp.2: Folgende Projektdaten können aus wettbewerblichen Gründen nicht veröffentlicht und zur Nachnutzung bereitgestellt werden …*