

Sammlungsstrategie

Einleitung

Das Forschungsdatenzentrum IANUS, das vom Deutschen Archäologischen Institut für die nationale Fachgemeinschaft koordiniert und betrieben wird, baut als seine zentrale Dienstleistung ein fachwissenschaftliches Datenarchiv mit zugehörigem Nachweiskatalog auf. Der primäre Zweck dieser Aktivitäten ist es, eine professionelle langfristige Archivierung von Primärdaten aus allen Bereichen der altertumswissenschaftlichen Forschung zu gewährleisten sowie diese als qualitative und gut dokumentierte Datensätze für eine wissenschaftliche Nachnutzung zu veröffentlichen. Die Regeln der Deutschen Forschungsgemeinschaft (DFG) zur Sicherung guter wissenschaftlicher Praxis (1998), deren Ergänzungen (2013) und Leitlinien zum Umgang mit Forschungsdaten (2015) bilden dabei die Grundlage des Handelns von IANUS.

Dieses Dokument beschreibt den Umfang und die Art der Forschungsdaten, die IANUS für die genannten beiden Ziele sammelt und für die es bereit ist Verantwortung zu übernehmen. Wichtige Ergänzungen zu dieser *Sammlungsstrategie* bilden folgende Dokumente, die allerdings noch erstellt werden müssen:

- Erhaltungsstrategie, in der die Maßnahmen zur langfristigen Aufbereitung, Sicherung und Erhaltung der digitalen Objekte definiert werden
- Nutzungsbedingungen, welche die Zugriffsmöglichkeiten auf die digitalen Objekte beschreiben und die Bedingungen für eine Nachnutzung festlegen
- Leitlinien des Archivs, in dem die Verwaltung, Speicherung und Verarbeitung der digitalen Datensammlungen beschrieben werden

Altertumswissenschaftliche Forschungsdaten

Definition Forschungsdaten

Gegenstand der Bemühungen von IANUS sind digitale Forschungsdaten aus den Archäologien und Altertumswissenschaften. Der Begriff Forschungsdaten wird dabei in einem weiten Sinne verstanden, der alle Daten umschließt, die für die Forschung in diesen Disziplinen interessant sein können. Danach sind sowohl Primärdaten relevant, die zur gezielten Beantwortung wissenschaftlicher Fragen z. B. auf Ausgrabungen als Fotos und Zeichnungen, bei chemisch-physikalischen Analysen als Messreihen und Graphen oder bei Editionen in Form von Textannotationen erhoben werden, als auch solche, die z. B. als Lehrmaterialien zu Ausbildungszwecken, im Rahmen von Verwaltungs- und Inventarisationsvorgängen in Denkmalfachbehörden und Museen oder bei Digitalisierungsmaßnahmen erstellt werden.

Die Daten haben dabei im Allgemeinen eine deutlich längere Lebensdauer als die Projekte und Systeme, in denen sie entstanden sind. Ein großer Teil von ihnen bleibt unveröffentlicht oder existiert nur auf lokalen, nicht allgemein zugänglichen Speichermedien und wird daher häufig nur unzureichend oder lediglich als Auswahl in den wissenschaftlichen Diskurs eingebunden. Oftmals geraten diese digitalen Grundlagen in Vergessenheit, werden ganz oder teilweise gelöscht, können aufgrund technischer Weiterentwicklungen nicht mehr gelesen werden oder bleiben mangels ausreichender Dokumentation unverständlich. In der Konsequenz stehen viele der einmal gesammelten Informationen künftigen Generationen nicht mehr zur Verfügung.

Aus diesem Grund verfolgt IANUS zwei Sammlungsschwerpunkte, die sich auf unveröffentlichte wissenschaftliche Informationen beziehen, da diese in der Regel dem Lebenszyklus digitaler Forschungsdaten entzogen sind und in besonderer Weise für die Zukunft gefährdet sind:

- unbearbeitete, originale und oftmals einmalige (Primär-)Daten
- aus diesen abgeleitete und prozessierte (Sekundär-)Daten mit unterschiedlichen Arbeitsständen.

Darüber hinaus werden aber auch die digitalen Vorlagen zu analog gedruckten oder online publizierten Ergebnissen berücksichtigt und in das Archiv aufgenommen.

Arten von Forschungsdaten

Aktuelle Projektdaten

Anders zu bewerten sind Daten aus aktuellen oder erst kürzlich abgeschlossenen Projekten, bei denen Verantwortliche und/oder Datenbeauftragte erreichbar sowie Gelder für die Kuratierung und Archivierung vorhanden sind. Im Idealfall, d. h. bei einer frühzeitigen Kontaktaufnahme mit IANUS, kann bereits während der Durchführung eines Projektes eine Unterstützung bei der Erhebung, Verarbeitung und Analyse der Daten erfolgen und vorbereitende Maßnahmen zur Nachnutzbarkeit und Archivierbarkeit von Daten begleitet werden. Ein weiterer Vorteil bei aktuellen bzw. kürzlich abgeschlossenen Projektdaten ist, dass Personen existieren, welche den Kontext der Datenerhebung kennen, bei der Dokumentation und Erschließung der Daten mit Metadaten mitwirken können und potentiell für Rückfragen zur Verfügung stehen.

Altdaten

In vielen Einrichtungen gibt es erhebliche Mengen an sogenannten Altdaten oder verwaisten Daten. Darunter werden Datenbestände aus abgeschlossenen Projekten verstanden, für die keine regulären Finanzmittel mehr vorhanden, keine Personen aktuell tätig und Autoren als Kenner der Daten häufig nicht mehr erreichbar sind. Als Ansprechpersonen fungieren oftmals nur Vertreter von Institutionen, an denen ein Projekt durchgeführt wurde, die aber nicht persönlich bei der Genese und Verarbeitung dieser Datenbestände beteiligt waren und diese daher in der Regel nur unzureichend kennen. Manchmal liegen die Projektlaufzeiten bereits mehrere Jahre zurück, so dass zusätzlich die Gefahr besteht, dass die Dateien technisch nicht mehr oder nur mit verhältnismäßig großem Aufwand gelesen werden können und darüber hinaus erforderliche Metainformationen zum Verständnis fehlen bzw. über die Zeit verloren gegangen sind.

Dennoch sind auch diese Daten für die zukünftige Forschung oft von großer Bedeutung und es besteht die akute Bedrohung, dass sie dem Erkenntnisprozess in den Archäologien und den Altertumswissenschaften dauerhaft entzogen werden. Es ist daher das langfristige Ziel von IANUS, möglichst viele Altdaten in das Langzeitarchiv zu übernehmen. Jeder Bestand ist dabei individuell zu bewerten und wird häufig einen überdurchschnittlich hohen Aufwand an technischer Aufbereitung und inhaltlicher Erschließung erfordern. Die konkret benötigte Zeit hängt jeweils von dem Zustand und der Qualität der Altdaten ab. Ganz grob lassen sich unterschiedliche Klassen von Altdaten unterscheiden:

- analoge Daten, die nicht in digitalisierter Form vorliegen oder
- digitalisierte Daten, die sich auf veralteten Speichermedien befinden und/oder
- digitalisierte Daten, die in nicht mehr unterstützten Dateiformaten vorliegen und/oder
- digitalisierte Daten, für die keine bzw. eine nicht ausreichende Dokumentation existiert oder deren Metadatenangaben technisch nicht unterstützt werden und/oder
- digitalisierte Daten, deren Metadaten in einem geeigneten Zielformat existieren

Kontinuierliche Daten

Eine dritte Gruppe von Daten, die für IANUS relevant sind, sind solche aus laufenden, mehrjährigen Projekten und Daueraufgaben, für die kein zeitlich verbindliches Ende festgelegt werden kann. In diesen Fällen, z. B. Corpora-Projekte oder Langzeitgrabungen, kann es sinnvoll sein, zu einem bestimmten Zeitpunkt einen Teil der Daten aus den Arbeitssystemen zu extrahieren und zu archivieren. Ähnliches gilt auch für laufende Fachsysteme, die kontinuierlich gepflegt und aktualisiert werden und für die es keinen definierten Abschluss wie bei Projekten mit vorgegebenen Laufzeiten gibt. Bei der Archivierung und Bereitstellung dieser Daten werden dann zusätzliche Informationen wie Versionsangaben und Verweise auf zugehörige, ältere Datensammlungen erhoben und Nutzern entsprechend angezeigt.

Analoge Archivmaterialien

Es wird explizit darauf hingewiesen, dass IANUS nur digital vorliegende Informationen archiviert und bereitstellt. Es besitzt keine Expertise und Infrastruktur für die langfristige Aufbewahrung papiergebundener Dokumentationen und analoger Archivmaterialien, weshalb derartige Datenbestände nicht angenommen werden. IANUS kann aber bei der Digitalisierung analoger Materialien beraten und bei der Beantragung von Projektmitteln zu diesem Zweck unterstützen. Auch wenn IANUS keine analogen oder papiergebundenen Forschungsdokumente aufbewahrt und aufbereitet, ist es dennoch an der Erfassung und Verbreitung von strukturierten Informationen über vorhandene Archivbestände interessiert und sammelt entsprechende Metadaten im Rahmen seines Nachweiskataloges.

Personenbezogene Daten

Alturtumswissenschaftliche Daten können sensible oder vertrauliche Informationen enthalten, die sich auf einzelne, identifizierbare Individuen beziehen und gleichzeitig wichtige inhaltliche oder historiographische Angaben zum Verständnis einer Datensammlung liefern. In einem weiteren Sinne können sie auch eine Bedeutung für die Wissenschaftsgeschichte einer alturtumswissenschaftlichen Fachdisziplin und ihrer Akteure besitzen. Einerseits ist IANUS bestrebt, auch solche Informationen zu erhalten und für die weitere Forschung zugänglich zu machen; andererseits erkennt es, dass dadurch in besonderem Maße Fragen des Datenschutzes, zu Persönlichkeitsrechten und zur Vertraulichkeit aufgeworfen werden, die in eigenen Richtlinien zur Übernahme von sensiblen Daten thematisiert werden.

Thematische Ausrichtung

Ob ein digitaler Datenbestand geeignet ist für die Aufbereitung, Archivierung und Bereitstellung durch IANUS und in dessen Sammlungsprofil passt, hängt von der thematischen Relevanz und von formalen Kriterien ab. Dabei erfolgt jeweils eine Einzelfallentscheidung über die Annahme oder Ablehnung von angebotenen Datensammlungen. Ganz generell umfasst das Sammlungsprofil von IANUS alle Arten von archäologischen und alturtumswissenschaftlichen Forschungsdaten, unabhängig davon, um welche Datentypen es sich handelt, wer der Datengeber ist und mit welchen Forschungsmethoden und in welchen Fachdisziplinen die Daten erzeugt wurden.

Fachrichtungen

Die Dienstleistung der Langzeitarchivierung wird allen Mitgliedern aus Fachrichtungen angeboten, die sich weltweit mit der Erforschung antiker Kulturen beschäftigen und im weitesten Sinne den Alturtumswissenschaften zuzurechnen sind. Einen darüber hinaus gehenden fachlichen Schwerpunkt gibt es nicht, so dass Daten zu Kulturen des vorderen Orients, des Mittelmeerraumes, Afrikas und Mitteleuropas genauso willkommen sind wie zu geographisch weiter entfernte Kulturen, etwa im eurasischen Raum, in Mittel- und Südamerika oder Ozeanien. Entscheidend ist lediglich die Tatsache, dass ausreichende fachspezifische Kompetenzen – entweder innerhalb von IANUS oder durch Hinzuziehung externer Experten – vorhanden sein müssen, um eine adäquate, d. h. inhaltlich qualitative Datenaufbereitung durchführen zu können.

Die Grenzen zu anderen Fachdomänen sind dabei fließend, z. B. in zeitlicher Hinsicht zu der frühen Menschheitsgeschichte, dem Mittelalter und der Neuzeit oder in methodischer Hinsicht zur Sinologie, Afrikanistik, Ethnologie und Kunstgeschichte. Ein besonderes Augenmerk wird auf archäologisch arbeitende Fachdisziplinen gelegt, da bei diesen in großem Umfang ganz neue Primärdaten erzeugt werden, die aufgrund der mangelnden Reproduzierbarkeit einzigartig und daher in ihrer langfristigen Verfügbarkeit besonders gefährdet sind.

In diesem Sinne richtet sich IANUS insbesondere an die folgenden Fachrichtungen.

Fachrichtungen mit einem Fokus auf materielle Hinterlassenschaften:

- Klassische Archäologie
- Provinzialrömische Archäologie
- Urgeschichte / Prähistorische Archäologie
- Frühgeschichte
- Mittelalterarchäologie
- Vorderasiatische Archäologie
- Orientalische Archäologie
- (Früh-)Christliche und byzantinische
- Archäologie
- Biblische Archäologie
- Islamische Archäologie
- Ägyptische Archäologie
- Altamerikanistik
- Afrikanische Archäologie
- Montanarchäologie

Fachrichtungen zur Analyse von Sprache und Literatur:

- Lateinische Philologie
- Griechische Philologie
- Altorientalistik / Assyriologie
- Byzantinistik
- Ägyptische Philologie
- Koptologie
- Hethitologie

Fachrichtungen zur Auswertung und Interpretation historischer Quellen und Schriftzeugnisse:

- Alte Geschichte
- Mittelalterliche Geschichte
- Papyrologie
- Epigraphik
- Numismatik

Weitere Disziplinen, die für die Erforschung vergangener Kulturen relevant sind:

- Geoarchäologie
- Archäometrie
- Dendrochronologie
- (Antike) Architekturgeschichte /
Bauforschung
- Anthropologie
- Paläobotanik
- Archäozoologie
- Paläogenetik
- Geophysik

Institutionen

Wie in vielen anderen Disziplinen auch, werden in den deutschen Altertumswissenschaften Forschungsdaten nicht nur unmittelbar an Universitäten oder außeruniversitären Forschungseinrichtungen erzeugt, sondern ebenso in verschiedenen privaten und staatlichen Einrichtungen. Aufgrund des öffentlichen Interesses an der Thematik und ihrer Bedeutung für den Erhalt des kulturellen Erbes existieren Museen und Archive, die neben der Aufbewahrung und Präsentation von Objekten auch die Aufgabe haben, Inhalte der Archäologie und der Altertumswissenschaften der Bevölkerung zu vermitteln, z.B. im Rahmen von Ausstellungen, Publikationen oder Führungen. Daneben existieren Landesdenkmalämter und weitere Fachbehörden, die in einem festgelegten geographischen Raum die gesetzlichen Aufgaben der Denkmalpflege und des Kulturgüterschutzes wahrnehmen und dabei teilweise auch einen expliziten Forschungsauftrag besitzen. Je nach den gesetzlichen Rahmenbedingungen können die Durchführung und Dokumentation von Ausgrabungen, geophysikalischen Prospektionen, Materialanalysen und Visualisierungen an kommerzielle Unternehmen vergeben werden, die forschungsnahe Dienstleistungen erbringen und wiederum eine eigenständige Gruppe bilden.

Für alle Akteure aus diesen Institutionen, d.h.

- Außeruniversitäre
Forschungseinrichtungen
- Universitäten und Hochschulen
- Museen, Sammlungen und Archive
- Denkmalfachbehörden
- Kommerzielle Dienstleister
- Fachverbände, Vereine,
Arbeitsgemeinschaften

ist IANUS bereit, in einem nicht-exklusiven Vertragsverhältnis die Langzeitarchivierung zu übernehmen, unabhängig davon, ob es sich im konkreten Fall um Individuen, Projektgruppen, Unternehmen oder Institutionen als Rechteinhaber an den jeweiligen Forschungsdaten handelt.

Ort und Zeit

IANUS bereitet auf, archiviert und veröffentlicht Daten unabhängig von ihrem geographischen Bezug. Auch wenn IANUS als nationales Forschungsdatenzentrum einen besonderen Schwerpunkt auf Daten legt, die sich unmittelbar auf Deutschland beziehen und für diese aufgrund seines Auftrages und seiner Finanzierung eine besondere Verantwortung übernimmt, ist dies nicht als Einschränkung seines Sammlungsprofils zu verstehen. Es werden grundsätzlich alle Daten zu vergangenen Kulturen und Forschungsaktivitäten in der gesamten Welt angenommen, sofern die formalen Kriterien (s. u.) erfüllt werden. Insbesondere bei Datenbeständen, die zu im Ausland durchgeführten Projekten gehören, wird auf eine mehrsprachige Dokumentation Wert gelegt, damit die Daten mit Infrastrukturen in den Herkunftsländern ausgetauscht werden können und so auch im jeweiligen Ausland ein adäquater Zugang gewährleistet werden kann.

Wie bei der geographischen Relevanz gibt es auch für die Zeitstellungen der untersuchten Kulturen keine Begrenzung bzw. wird diese über die Forschungsgegenstände der genannten Fachdisziplinen definiert. Die Daten können sich also ebenso auf die früheste Menschheitsgeschichte beziehen wie auf Gesellschaften der jüngsten Vergangenheit. Auch hierbei gilt der Grundsatz, dass IANUS zwar bereit ist, Daten jeglicher Zeitstufe anzunehmen, im Einzelfall aber immer auch die Möglichkeiten anderer fachspezifischer Archive und Repositorien prüft.

Methoden

Da unter dem Begriff Altertumswissenschaften ganz unterschiedliche Disziplinen aus den Geistes- und Naturwissenschaften zusammengefasst werden, weisen die Daten, die für die Untersuchung vergangener Kulturen relevant sind, eine hohe fachlich-methodische Heterogenität auf. Typische Arbeitsweisen sind z.B.:

- archäologische Ausgrabungen
- Oberflächenbegehungen (Survey)
- geophysikalische Prospektionen
- Entnahme geologischer Bohrkerne
- architektonische Bauaufnahmen und Rekonstruktionen
- naturwissenschaftliche Untersuchungen anorganischer Materialien
- Analyse von Knochen- und Pflanzenresten
- Vermessung und Geodäsie
- Anfertigung von Laserscans und 3D-Dokumentation
- Analyse geographischer Räume und Landschaftsmodellierung
- Auswertung von Luftbildern und Satellitenaufnahmen
- syntaktische und semantische Quellenkritik
- Annotation von Texten

Datentypen

Genauso vielfältig wie die Fachrichtungen, Institutionen und Methoden sind in den Altertumswissenschaften auch die verwendeten Datenformate. Auch wenn IANUS zur Aufbereitung und Archivierung nicht alle existierenden Dateiformate annehmen wird (siehe ausführlich zum Konzept der akzeptierten und präferierten Dateiformate die *Erhaltungsstrategie* sowie die Hinweise zur „Archivierung bei IANUS“ in den [IT-Empfehlungen](#)), so werden verschiedene Datentypen aufgrund ihrer Relevanz für die Altertumswissenschaften angenommen, aufbereitet und archiviert. Die wichtigsten sind:

- Vektorgrafiken / Zeichnungen / CAD
- Rastergrafiken, insbesondere digitale Fotos
- freie Texte und beschreibende Dokumente
- strukturierte Daten wie Tabellen, Listen, Kataloge
- statistische Daten
- Datenbanken
- Koordinaten, topographische Karten, Vermessungsinformationen und raumbezogene Daten, insbesondere aus geographischen Informationssystemen
- Luftbildaufnahmen und Daten der Fernerkundung
- geophysikalische Messdaten
- naturwissenschaftliche Messreihen
- Punktwolken von Laserscannern
- Rekonstruktionen / 3D-Daten / Virtual Reality
- Mark-Up-Dateien und Webseiten
- elektronische Versionen von Zeitschriften, Journalen und Büchern
- unpublizierte Berichte (sog. „grey literature“) und Qualifikationsarbeiten
- Audio und Video

Die Aufbereitung und Archivierung von eigenständigen Software-Anwendungen wird dagegen nicht von IANUS gewährleistet, da ihre Nutzbarkeit häufig von bestimmten Hardwareumgebungen abhängt, die oftmals nur für einen bestimmten, überschaubaren Zeitraum relevant sind und nur mit großem programmiertechnischem Aufwand in neue Systeme migriert werden können.

Archivierung von Daten – Nachweis von Daten

Im Kontext der Sammlungsstrategie von IANUS sind grundsätzlich zwei Arten von Daten zu unterscheiden:

- Sammlungen von forschungsrelevanten digitalen Daten, die IANUS mit dem Ziel der Aufbereitung, Archivierung und Bereitstellung anvertraut werden
- Metadaten über analoge und digitale Fachressourcen, die bei externen Institutionen aufbewahrt werden

Beide Arten werden gemeinsam in einem Online-Portal zur Verfügung gestellt, mit dem Unterschied, dass im ersten Fall je nach den vom Datengeber festgelegten Zugriffsbeschränkungen und nach Zustimmung zu den Nutzungsbedingungen von IANUS diese Dateien direkt zur Nachnutzung heruntergeladen werden können. Im zweiten Fall werden lediglich die Metadateninformationen angezeigt, der eigentliche Zugriff auf die Ressourcen erfolgt aber über die datenhaltende externe Einrichtung und deren Systeme.

Grundsätzlich ist IANUS bestrebt, alle zur Verfügung stehenden Daten und Metadaten so offen und frei wie möglich zu präsentieren. Dennoch können diese sensible und schützenswürdige Informationen enthalten (etwa persönlich Angaben zu Einzelpersonen; Lokalisierungen von schutzwürdigen Untersuchungsflächen; kurz vor der Publikation stehende Ergebnisse), für die gestaffelte Zugriffsrechte erforderlich sein können. Für bestimmte Nutzergruppen können daher einzelne Metadaten und digitale Objekte ein- oder ausgeblendet oder in Ihrer Genauigkeit angepasst werden. Letzteres ist insbesondere bei Geoinformationen von hoher Relevanz: Bei anonymen, nicht registrierten Benutzern wird deshalb nur die ungefähre Lage angezeigt und detaillierte Ortsbezeichnungen ausgeblendet werden, während Benutzer, die z. B. durch eine Anmeldung als Mitarbeiter einer bestimmten Einrichtung authentifiziert sind, exakte Koordinaten einsehen und nutzen dürfen. Welche Zugriffsgruppen existieren und welche Einschränkungen oder Befugnisse diese haben sollen, wird durch die Bereitsteller der Daten und Metadaten in Zusammenarbeit mit IANUS festgelegt.

Archivierung von Datensammlungen

Grundlegend für das Verständnis von Forschungsdaten bei IANUS ist das Konzept der Datensammlung, die aus einer Summe von inhaltlich, rechtlich und logisch zusammengehörenden digitalen Objekten und den beschreibenden Metainformationen besteht. Eine Datensammlung bildet somit eine fachlich relevante Einheit, die als solche aufbereitet und veröffentlicht wird. Dabei ist es irrelevant, ob es sich um Altdaten, aktuelle Projektdaten oder kontinuierliche Daten handelt.

Eine Datensammlung wird von IANUS archiviert mit dem Ziel, ihre wissenschaftlichen Inhalte so lange wie möglich zu erhalten und aufzubewahren. Individuen, Projekte, Institutionen und Unternehmen können – z. B. am Ende einer Förderung oder beim Projektabschluss – ihre Daten mit dem Auftrag zur Langzeitverfügbarmachung bei IANUS abgeben, wo sie nach definierten Workflows und Standards formal überprüft, semantisch vervollständigt, technisch aufbereitet und ausführlich dokumentiert werden, um diese dann langfristig zu archivieren und bereitzustellen.

Für den Eingang einer neuen Datensammlung in den Archivbestand von IANUS ist die Unterzeichnung eines Datenübergabevertrages erforderlich. Eine Bedingung zur Annahme von angebotenen Daten ist, dass IANUS während der vertraglichen Zusammenarbeit möglichst weitgehende Rechte für diese eingeräumt werden, damit IANUS diese bearbeiten also verändern (Formatmigrationen) darf, den Zugriff gewährleisten und eine Verbreitung ermöglichen kann. Die vertragliche Vereinbarung mit IANUS ist dabei nicht-exklusiv, so dass der Rechteinhaber einerseits weiterhin über seine Daten unabhängig von IANUS verfügen kann und andererseits seine Daten auch an Dritte weitergeben kann. Datensammlungen, für die sehr hohe restriktive Auflagen in Bezug auf Zugänglichkeit, Nachnutzung und Verbreitung eingefordert werden, werden nur in besonderen Fällen angenommen.

Den Umfang und den Inhalt einer Datensammlung legt zunächst der Datengeber selbständig fest und wählt diejenigen Forschungsdaten aus, die er für eine Langzeitarchivierung und Nachnutzung geeignet hält. Die Frage, ob der gesamte Datenbestand eines Projektes oder einer Institution eine einzige oder eher mehrere separate Datensammlungen darstellt, sollte abschließend nach Rücksprache mit IANUS entschieden werden. Häufig ist es sinnvoll, sehr große und heterogene Datenbestände in mehrere kleinere Datensammlungen aufzuteilen und diese durch Referenzen miteinander in Beziehung zu setzen.

Damit eine digitale Datensammlung archiviert werden kann, ist es notwendig, dass ihr Inhalt statisch ist, d. h. sie umfasst eine in sich abgeschlossene Menge von Fachinformationen, die bei der Übergabe an IANUS in einem finalen Zustand vorliegen und sich nicht mehr ändern – ähnlich wie dies auch bei der Übergabe von analogen Dokumenten an ein klassisches Archiv erforderlich ist. Der Grund hierfür liegt darin, dass Daten, die einmal den Prozess der Aufbereitung und Formatmigration durchlaufen haben und in ein Offline-Storage-System zur Archivierung und Bitstream-Speicherung überführt wurden, im Nachhinein nur mit erheblichem Aufwand geändert oder gelöscht werden können. Es können zwar Ergänzungen und Nachlieferungen zu bestehenden Archivbeständen vorgenommen werden, doch werden diese als separate, eigenständige Informationsobjekte angesehen. Dies bedeutet im Einzelfall, dass mit den an IANUS übertragenen Daten zwar noch weitergearbeitet werden kann (z. B. zur Vorbereitung von Publikationen oder zur Durchführung neuer Analysen), die eigentlichen (Primär-)Daten dadurch aber nicht mehr verändert werden. Sofern neue zugehörige (Sekundär-)Daten entstehen und diese ebenfalls archiviert werden sollen, werden sie nicht der ursprünglichen Datensammlung hinzugefügt, sondern als neuer Datenbestand in das Archiv aufgenommen. Insofern sollte nicht nur der Umfang einer Datensammlung, sondern auch der geeignete Zeitpunkt der Übergabe mit IANUS abgestimmt werden.

Löschung von Archivbeständen

Im Allgemeinen geht IANUS davon aus, dass Datensammlungen, die einmal aufbereitet und archiviert wurden, nicht mehr gelöscht werden sollen. In seltenen Fällen jedoch, in denen das Entfernen einer Datensammlung aus dem Archivbestand von IANUS als sinnvolle oder notwendige Maßnahme erachtet wird, wird eine endgültige Entscheidung über einen solchen Schritt erst nach Rücksprache mit dem ursprünglichen Datengeber bzw. dem aktuellen Rechteinhaber gefällt. Sofern erwünscht und möglich werden die originalen und archivierten Daten dann an den Rechteinhaber zurückgegeben und/oder es wird ein Transfer in ein geeignetes alternatives Datenarchiv unterstützt.

Nachweis von Fachressourcen

Neben der Archivierung und Bereitstellung von Forschungsdaten bietet IANUS im Rahmen seines Online-Portals auch einen Nachweiskatalog an, in dem Fachinformationen über Datenbestände und Ressourcen gesammelt werden, die von unterschiedlichen wissenschaftlichen Akteuren verwaltet, betrieben und gepflegt werden. Die eigentliche analoge oder digitale Ressource wird also außerhalb von IANUS aufbewahrt, IANUS wird aber auch auf sogenannte externe Ressourcen verweisen, dazu werden auf dem

Datenportal die Metadaten auf Sammlungsebene dargestellt und auf die Ressource (analog oder digital) verwiesen.

Die Idee dahinter ist, möglichst viele altertumswissenschaftliche Angaben zu Fundstellen und Monumenten, wissenschaftlichen Projekten, Akteuren und Ressourcen (d.h. analoge und digitale Sammlungen von Dokumenten und Daten) zu bündeln und gemeinsam zu präsentieren. Hierbei steht die Zusammenführung, Integration und Anreicherung von fachspezifischen und heterogen modellierten Metadaten im Vordergrund, um trotz unterschiedlicher Akteure und Systeme einen einheitlichen Zugriff und eine vergleichbare Datenqualität zu gewährleisten.

Die zentralen Kategorien des Nachweiskataloges sind:

- *Akteure*, also Institutionen, Unternehmen, Gruppen oder Personen, die Aktivitäten und Projekte durchführen, für archäologische Fundstellen verantwortlich sind oder Ressourcen besitzen.
- *Archäologische Fundstellen bzw. Archäologieflächen* wie sie im ADEX-Standard des Verbandes der Landesarchäologen definiert werden, also Flächen mit einem im Gelände abgrenzbaren oder lokalisierbaren Bereich, an der mindestens ein archäologisch qualifiziertes bzw. relevantes Ergebnis vorliegt/vorgelegen hat oder vermutet wird. Zu dieser Kategorie gehören im weitesten Sinne auch Monumente, Bauwerke und andere Denkmäler.
- *Aktivitäten* wie wissenschaftliche Untersuchungen, Forschungsprojekte und andere notwendige Maßnahmen, für die eindeutig zuweisbare Akteure, eine feste Zeitdauer (Anfang und Ende einer Aktivität), eine spezifische Aufgaben- bzw. Fragestellung und ein geographischer Raumbezug bekannt sind.
- *Ressourcen* können Sammlungen analoger Dokumente (z. B. Grabungsunterlagen) und digitaler Daten jeglicher Art umfassen, die im Rahmen von Projekten oder von Daueraufgaben entstehen. Wichtig ist, dass eine Ressource immer mit einem Eigentümer oder einer Ansprechperson verbunden werden kann. In dieser Kategorie fallen auch die digitalen Bestände von IANUS.

Durch das Angebot und die Pflege eines zentralen Nachweiskataloges soll für WissenschaftlerInnen die Recherche nach fachspezifischen Informationen, die über Deutschland mit seinen verschiedenen Bundesländern verteilt sind, erleichtert werden. In technischer Hinsicht dient dieser Katalog auch zum Suchen und Finden der Inhalte in dem Archivbestand von IANUS. Konzeptionelles Vorbild sind verschiedene nationale Portale anderer Ländern (z.B. Archsearch¹ des ADS in Großbritannien, fund og fortidsminder² in Dänemark, Riksantikvarieämbetet³ in Schweden), in denen Metadaten zu Monumenten und Fundstellen zusammengeführt werden.

Der Nachweiskatalog wird vor allem folgende Funktionalitäten beinhalten:

- Verschiedene Möglichkeiten zur Navigation durch die Daten: Über die Eingabe von Suchbegriffen; mittels Browsing durch festgelegte Schlagworte und Kategorien; Kartenbasiert mit Möglichkeiten zum Zoomen und Ziehen eines Auswahl-Polygons
- Anzeige von Suchergebnissen mit Metadaten, Kartendarstellung und Verlinkungen zu externen Systemen für weitere Informationen
- Export und Download von Ergebnismengen, z. B. als Listen oder Karten zur individuellen Weiterverarbeitung auf lokalen Systemen

Die jeweiligen datenhaltenden Einrichtungen, z.B. Landesdenkmalämter, Museen, Archive oder Forschungsinstitute behalten dabei alle Rechte an und die Zuständigkeit zur Bereitstellung und Pflege der Originaldaten. Sie sind auch für die Qualität und Aktualität der Metadaten verantwortlich, die im Nachweiskatalog angezeigt werden.

¹ <http://archaeologydataservice.ac.uk/archsearch/>

² <http://www.kulturarv.dk/fundogfortidsminder/>

³ <http://www.fmis.raa.se/cocoon/fornsok/search.html>

Die zentralen Eigenschaften des Nachweiskataloges sind:

- Alle Einzelnachweise für Informationen, die sich nicht auf archivierte Datensammlungen bei IANUS beziehen, werden nur als Metadatensatz vorgehalten.
- Einzelne, verpflichtende Metadaten weisen explizit auf die datenhaltende Einrichtung hin.
- Sofern es sich bei den nachgewiesenen Daten um digitale und online verfügbare Ressourcen handelt, werden entsprechende Verlinkungen auf die Ursprungssysteme angezeigt.

Soweit fachlich sinnvoll und technisch möglich wird IANUS aktiv versuchen, mit datenhaltenden Einrichtungen und Infrastrukturen eine Kooperation über den Austausch von und Zugriff auf Metadaten zu erzielen, um den fachwissenschaftlichen Wert des Nachweiskataloges zu erhöhen.

Qualitätssicherung

Um sowohl für Datengeber als auch für Datennachnutzer einerseits einen größtmöglichen Mehrwert von archivierten und bereitgestellten Daten schaffen zu können und andererseits ein ausgewogenes Verhältnis von konkretem Archivierungsaufwand und wissenschaftlichem Nachnutzungspotential zu erlangen, spielen für die Übernahme von Datenbeständen durch IANUS formale, technische und inhaltliche Aspekte eine wichtige Rolle. IANUS wird individuell prüfen, inwieweit eine zur Archivierung angebotene Datensammlung die im folgenden beschriebenen Kriterien erfüllt oder nicht und fallweise entscheiden, ob und unter welchen Bedingungen eine Übernahme in die Archivbestände von IANUS möglich ist, ob Daten abgelehnt werden müssen oder ob die Übergabe an ein anderes, besser geeignetes Datenzentrum sinnvoll ist.

Formale Kriterien

Damit IANUS die volle langfristige Verantwortung für anvertraute Datenbestände übernehmen kann, sind einige formale Kriterien zu beachten. Diese sollen vor allem dem Umstand Rechnung tragen, dass IANUS die archivierten Daten technisch und inhaltlich verstehen kann und rechtlich wie finanziell dazu in der Lage ist. Im Allgemeinen ist IANUS bereit, diejenigen digitalen Daten anzunehmen, die für die Beantwortung kulturgeschichtlicher Fragen über antike Zivilisationen relevant sind und alle nachfolgend genannten Kriterien erfüllen.

- Die Daten sind aus fachlicher wie technischer Sicht abgeschlossen und ändern sich nicht mehr.
- Die Daten passen thematisch in das Sammlungsprofil von IANUS.
- Die Daten sind in deutscher und/oder englischer Sprache dokumentiert.
- Die Daten sind in rechtlicher Hinsicht unkritisch, d. h. alle Fragen zu Rechteinhabern und zu Nutzungs- und Verwertungsrechten sind geklärt.
- Die Daten wurden
 - entweder unter Beteiligung von deutschen Forschern erzeugt,
 - oder durch deutsche Mittelgeber (mit)finanziert
 - oder beziehen sich auf Deutschland
- Die Daten erfüllen die von IANUS in den IT-Empfehlungen veröffentlichten technischen und dokumentarischen Vorgaben und sind dementsprechend aufbereitet.

Inhaltliche und technische Kriterien

Bei der inhaltlichen Bewertung einer Datensammlung und bei der Klärung der Frage, inwieweit eine angebotene Datensammlung zum Sammlungsprofil von IANUS passt, kommen vor allem folgende Aspekte zum Tragen:

- Die wissenschaftliche Relevanz und das Potential für eine künftige Nachnutzung
- Die generelle Eignung eines Datenbestandes für die dauerhafte Aufbereitung, Archivierung und Bereitstellung für Dritte
- Die Existenz oder das Fehlen alternativer Infrastrukturen zur Aufbewahrung von Daten

Während das erste Kriterium primär den Inhalt einer Datensammlung betrifft, konzentriert sich das zweite auf die Organisation der Daten und die Dateiformate, auf die mitgelieferte Dokumentation sowie auf rechtliche Fragen. Der dritte Aspekt dient vor allem dazu, einen möglicherweise doppelten Aufbereitungsaufwand zu vermeiden, ggf. Empfehlungen für Infrastrukturen mit spezifischeren Kenntnissen und Erfahrungen auszusprechen oder – besonders im Falle von Daten, die sich auf die antiken Kulturen außerhalb Deutschlands beziehen – Kontakt mit ausländischen Archiven aufzunehmen.

Die Evaluierung von angebotenen Datensammlungen nach diesen Aspekten erfolgt zu einem sehr frühen Zeitpunkt, um rechtzeitig abschätzen zu können, welche Maßnahmen für eine langfristige Erhaltung und Nachnutzung notwendig sein werden und welche Kosten voraussichtlich entstehen werden.

Es ist zu betonen, dass auch Datensammlungen, die nicht alle Kriterien befriedigend erfüllen, nicht notwendigerweise und automatisch abgelehnt werden. Insbesondere wenn sie einen hohen wissenschaftlichen Wert und ein hohes Nachnutzungspotential besitzen, werden stattdessen in der Dokumentation entsprechende Hinweise formuliert und künftige Nutzer auf mögliche Defizite und Einschränkungen explizit hingewiesen. Dies kann insbesondere bei Altdaten der Fall sein.

Wissenschaftliche Qualität

IANUS ist bestrebt, möglichst qualitätsvolle Daten zu übernehmen und bereitzustellen, d. h. die Daten entsprechen den üblichen Qualitätsstandards der jeweiligen altertumswissenschaftlichen Fachdisziplinen und sie erleichtern in hohem Maße die künftige Forschung. Dies gilt insbesondere für Daten, die als primäre Quellen nicht reproduzierbare Forschungstätigkeiten und Ergebnisse dokumentieren. Im Folgenden werden die Leitlinien beschrieben, die bei der Bewertung einer Datensammlung zum Tragen kommen. Dabei ist zu betonen, dass nicht die fachwissenschaftlichen Inhalte und Ergebnisse eines Projektes, sondern lediglich die Qualität der erzeugten digitalen Daten geprüft werden.

Nachnutzungspotential

In der Regel werden Ordner, Dateien und Systeme für einen spezifischen Zweck von Projekt-, Institutsmitarbeitern, Gruppen oder Individuen angelegt und sind in diesem konkreten Entstehungskontext verständlich und nachvollziehbar. Die Daten, die IANUS archiviert und bereitstellt, sollen dagegen für einen möglichst großen Nutzerkreis aus den genannten Fachgemeinschaften verstehbar und nutzbar sein. Es daher zu beurteilen, inwieweit eine Datensammlung prinzipiell für eine unbekannte künftige Nachnutzung geeignet ist oder durch entsprechende Aufbereitungsmaßnahmen in einen solchen Zustand überführt werden kann. Die Nutzbarkeit einer Datensammlung wird dabei ganz wesentlich von den fachspezifischen Anforderungen bestimmt. Da sich diese allerdings im Laufe der Zeit ändern können, wird IANUS regelmäßig überprüfen, inwieweit seine Bewertungskriterien den aktuellen Bedürfnissen und Kenntnissen der Nutzer gerecht werden.

Daneben hängt das Nachnutzungspotential auch von den verwendeten technischen Formaten und Systemen ab. Wenn z. B. ein Projekt vor allem proprietäre Software zur Erzeugung und Verwaltung seiner digitalen Daten verwendet und diese Software in der Fachgemeinschaft kaum genutzt wird, können die meisten Personen diese Daten nicht nutzen, so dass ohne entsprechende Gegenmaßnahmen das Nachnutzungspotential als sehr gering einzustufen ist.

Adäquate Dokumentation

Die Qualität einer Datensammlung hängt aus archivischer Sicht nicht allein von dessen wissenschaftlicher Qualität ab, sondern von der Art und dem Umfang der begleitenden notwendigen Dokumentation. Diese muss in angemessener Form sowohl den Inhalt beschreiben als auch technische Aspekte beleuchten. Aus einer guten Dokumentation sollte vor allem hervorgehen, in welchem Kontext und mit welcher Fragestellung Daten erzeugt wurden, wie diese vor der Übergabe an IANUS verwaltet und bearbeitet wurden und welche Beziehungen es zu Quellen und Informationen außerhalb der Datensammlung gibt. Auch fehlende oder unvollständige Informationen, die sich als Einschränkungen für die Nachnutzung auswirken können, müssen explizit benannt werden.

Archivierbarkeit

Die Archivierbarkeit einer Datensammlung hängt zunächst von den technischen Dateiformaten ab, in denen die Daten gespeichert sind. Wenn diese bereits bei der Übergabe der Daten an IANUS veraltet und überholt sind und auch von einem Datengeber selbst nicht in jüngere Formate konvertiert werden, also kaum oder nur mit erheblichem Aufwand in aktuelle archivtaugliche Formate überführt werden können, kann dies ein Grund für die Ablehnung von Daten sein. IANUS hat daher eine Liste von akzeptierten und präferierten Dateiformaten erarbeitet und in seinen IT-Empfehlungen publiziert. Über angebotene Formate, die dort nicht genannt werden, wird je nach Einzelfall entschieden.

Alternative Archivierung

Ein weiterer Aspekt ist die objektive Notwendigkeit, eine Datensammlung bei IANUS zu archivieren. Damit verbunden ist die Frage, ob sie bereits in einem anderen Repositorium oder digitalen Infrastruktur aufbereitet und archiviert wird oder ob ein akuter Handlungsbedarf besteht, um einen absehbaren Informationsverlust abzuwenden. Sofern eine Datensammlung bereits außerhalb von IANUS von einer Einrichtung professionell gespeichert und zur Verfügung gestellt wird, wird IANUS Kontakt mit dieser Einrichtung aufnehmen und zusammen mit dem Datengeber folgende Optionen abstimmen:

- Kann ein Austausch von Metadaten über die Datensammlung erfolgen, so dass die bisherige Einrichtung die Datensammlung auch weiter vorhält und für diese verantwortlich bleibt, sie aber auch über den Nachweiskatalog von IANUS auffindbar ist.
- Kann eine Kopie eines bereits existierenden Archivpaketes übernommen und in die Datenbestände von IANUS integriert werden, wenn dies für die Zielgruppen von IANUS von besonderer Relevanz ist, z. B. weil auf diese Weise eine fachspezifische Präsentation erfolgen kann.

Datentransfer

Datensammlungen für das Archiv

Der Transfer einer neuen Datensammlung eines Datengebers an IANUS kann auf unterschiedliche Weise erfolgen:

- Kleine Datenmengen können direkt mit Hilfe eines Upload-Formulars auf einen Server von IANUS kopiert werden. Dies bietet neben der zeitnahen Übertragung der Daten auch den Vorteil, dass eine automatisierte Überprüfung der akzeptierten Dateiformate erfolgen und dadurch schneller eine Rückmeldung gegeben werden kann.
- Für mittelgroße Datenmengen, deren Umfang immer noch eine Online-Übertragung zulässt, kommt ein ftp-Server bzw. ein eigener Cloud-Service zum Einsatz. Auch hier ist eine frühzeitige Prüfung von übertragenen Dateien möglich.
- Größere Datenmengen, für die eine Online-Übertragung nicht mehr in Frage kommen, müssen postalisch über externe Speichermedien (mobile Festplatten, USB-Sticks, SD-Karten etc.) an IANUS übermittelt werden, welche nach Überspielung auf die IANUS-Systeme an den Absender zurück gesandt werden.

Eine Übertragung von Daten ist bei allen drei Varianten erst nach einer Registrierung bzw. Anmeldung am Archiv-Management-System von IANUS und nach einer expliziten Aufforderung bzw. technischen Freischaltung durch einen Datenkurator von IANUS möglich.

Metadaten für den Nachweiskatalog

Die Art und Weise, wie neue Metadaten für den Nachweiskatalog von Dateneigentümern an IANUS übermittelt werden, kann technisch sehr unterschiedlich ausfallen. Das konkrete Verfahren hängt dabei ganz wesentlich von den jeweils involvierten Systemen ab. Denkbare Szenarien sind:

- eine direkte, bidirektionale Schnittstelle mit einem System. Dieser Weg ist technisch am aufwändigsten und kommt vor allem, aber nicht ausschließlich für die Systeme, die vom DAI zur

normierten Erfassung von Publikationen, Akteuren, Orten und Schlagworten, zur Anwendung.

- Übertragung mittels Web-Services und OAI-PMH-Schnittstellen, die z.B. im Rahmen der Geodateninfrastrukturen der einzelnen Bundesländer angeboten oder von größeren Institutionen (z. B. Akademien, universitären Datenzentren) zum Metadata-Harvesting bereitgestellt werden
- ein- oder mehrmalige (regelmäßige) Importe von offline zugesandten Daten (z.B. dateibasierte Exporte aus Fachsystemen)

Aktive und Passive Datenakquise

Bei der Umsetzung der Sammlungsstrategie von IANUS soll mittel- bis langfristig sowohl eine aktive als auch passive Akquise erfolgen, d. h. IANUS wird sowohl Datengeber, die sich auf Eigeninitiative bei IANUS melden begrüßen; als auch aktiv Eigentümer forschungsrelevanter Datensammlungen kontaktieren. Eine generelle Priorisierung zur Annahme oder Ablehnung von neuen Datensammlungen wird durch das Steuerungsgremium von IANUS festgelegt und hängt außerdem von den Anforderungen und Bedürfnissen der Fachwissenschaften ab. Ganz allgemein werden verschiedene Strategien zur Umsetzung der Sammlungsziele von IANUS angewandt:

Kooperationen mit datenhaltenden Einrichtungen

Sofern wissenschaftliche Einrichtungen über zentrale Bestände an altertumswissenschaftlichen Ressourcen verfügen, wird IANUS eine Kooperation mit diesen über den Austausch von Metadaten für seinen Nachweiskatalog anstreben und – sofern erwünscht – auch die Archivierung dieser Bestände übernehmen.

Vereinbarungen mit Drittmittelgebern

IANUS wird versuchen mit Drittmittelgebern, die für die altertumswissenschaftliche Forschung in Deutschland eine besondere Relevanz besitzen, Vereinbarungen dahingehend abzuschließen, dass geförderte Projekte dazu aufgefordert werden, nach ihrem Abschluss ihrer Daten an ein professionelles Archiv zu übergeben.

Übernahme von Datensammlungen individueller und institutioneller Datengeber

IANUS wird Datensammlungen von individuellen und institutionellen Datengebern annehmen, sofern diese die in diesem Dokument beschriebenen erforderlichen Kriterien erfüllen. Aktuelle und künftige Datengeber werden über die Verfahren zur Aufbereitung von Daten für die Langzeitarchivierung und Langzeitbereitstellung informiert und es werden ausreichende Hilfestellungen zur Vorbereitung von Datensammlungen vor dem Datentransfer zur Verfügung gestellt.

Projekte zur Erzeugung neuer Daten

IANUS wird sich an (Drittmittel-)Projekten beteiligen, die zur Erzeugung neuer Daten und Informationen dienen, sofern ein fachwissenschaftliches Defizit z. B. durch entsprechende Rückmeldungen der Nutzer oder durch Empfehlungen des Steuerungsgremiums erkennbar ist.