Automatisierte Integritätssicherung von wissenschaftlichen Primärdaten ab ihrer Erhebung

Jan Potthoff[1], Paul Christopher Johannes[2], Maximilian Stumpf[1]

[1]Karlsruher Institut für Technologie (KIT)
[2]Universität Kassel
jan.potthoff@kit.edu, paul.johannes@uni-kassel.de, maximilian.stumpf@kit.edu

Abstract: Um die Qualität der Forschung erhalten zu können, muss u. a. die Integrität der im Forschungsprozess entstandenen Daten nachweisbar sein. Primärdaten sind in der Forschung von besonderer Bedeutung, da auf ihnen das weitere Vorgehen, die Auswertung und die Ergebnisse beruhen. Eine frühzeitige Sicherung der Integrität ermöglicht es, Veränderungen an Daten aufzudecken. Um die Integrität frühzeitig zu sichern, können speziell dafür entwickelte Endgeräte eingesetzt werden. Diese stehen jedoch nur vereinzelt zur Verfügung. Durch eine generische Lösung kann dies aber auch für andere, bereits im Forschungsprozess eingesetzte Endgeräte verfügbar gemacht werden.

1 Einleitung

Daten fallen in allen Phasen des Forschungsprozesses in unterschiedlichen Formaten und Mengen an, also in Planung und Entwurf, Durchführung, Aufbereitung, Auswertung und Veröffentlichung und der abschließenden Archivierung [Ha11]. In der Planungs- und Entwurfsphase wird im Wesentlichen auf vorhandene Daten zurückgegriffen, um darauf basierend einen Entwurf für das weitere Vorgehen zu skizzieren. Der Entwurf wird in der Regel iterativ angepasst. Durch eine Versionierung und eine entsprechende Integritätssicherung kann dieser Prozessschritt nachvollziehbar gestaltet werden.

Auf die Nachvollziehbarkeit der Prozessschritte und darin erhaltenen Forschungsdaten kommt es schon aus Gründen der guten wissenschaftlichen Praxis an. Auf die Beweisbarkeit von diesen Aufzeichnungen kann es darüber hinaus in unterschiedlichen Szenarien, wie Patent- und Urheberrechtstreitigkeiten oder Zulassungs- und Kontrollverfahren ankommen. Für den einzelnen Wissenschaftler ist es insbesondere wichtig Vorwürfe wissenschaftlichen Fehlverhaltens ausräumen zu können. Die scientific community dagegen, will wissenschaftliches Fehlverhalten aufdecken und nachweisen können.

Wie zu diesem Zweck die Integrität und Authentizität der im Prozess entstandenen Daten im letzten Schritt, der Archivierung, gewährleistet werden kann, ist Ziel des Projekts "Beweissicheres elektronisches Laborbuch" (BeLab). Das Karlsruher Institut für Technologie (KIT), die Physikalisch-Technische Bundesanstalt (PTB) in Braunschweig und die Universität Kassel entwickeln dazu Konzepte, die eine beweiswerterhaltende Archivierung von Forschungsdaten gewährleisten sollen.

Die in der Durchführungsphase entstehenden Daten sind im Forschungsprozess von besonderer Bedeutung, da auf ihnen das weitere Vorgehen, die Auswertung und die Erkenntnisse und Ergebnisse beruhen. Des Weiteren können sie auch in anderen Forschungsvorhaben institutionsintern und eventuell extern genutzt werden. Die Bedeutung der Primärdaten fällt besonders ins Gewicht, wenn zugehörige Versuche nicht wiederholt und die Daten nur einmalig erzeugt werden können. Dies ist beispielsweise in der Klimaforschung bei der Aufzeichnung von Messdaten durch einen Wetterballon der Fall. Da Daten vorerst keiner Aufbereitung oder anderen Verarbeitungsschritten unterliegen, jedoch immer auf sie Bezug genommen werden kann, ist es sinnvoll die Primärdaten automatisiert in ihrer ursprünglichen Form zu sichern und die Integrität belegen zu können.

Im bisherigen Verlauf des BeLab-Projekts stand die beweissichere Archivierung von Forschungsdaten im Vordergrund, siehe dazu auch [Po12]. Um einen möglichst hohen Beweiswert zu erzielen, sollte die Absicherung der Integrität und Authentizität der Daten bereits in früheren Phasen des Forschungsprozesses erfolgen. Der Beitrag zeigt sich daraus ergebene Anforderungen und Lösungsmöglichkeiten u. a. durch einen generischen Ansatz.

2 Endgeräte im Forschungsprozess

Insbesondere in der experimentellen Forschung kommen Messgeräte und Analysegeräte (hier allgemein Endgeräte) zum Einsatz, deren Art, Nutzung und Output in Abhängigkeit zum Forschungsbereich und –vorhaben stehen. Des Weiteren sind Schnittstellen und genutzte Datenformate vom Hersteller des Endgeräts abhängig.

2.1 Schnittstellen zwischen Endgerät und Computer

Unter der Interprozesskommunikation wird im engeren Sinne der Austausch von Informationen zwischen zwei Prozessen, die keinen gemeinsamen Speicherbereich teilen, verstanden. Im weiteren Sinne ist die Interprozesskommunikation für getrennt laufende Systeme, wie beispielsweise eine Client- und Serveranwendung, von Bedeutung. So muss der Zugriff auf entsprechende Informationen geregelt sein, da kein gleichzeitiger Zugriff möglich ist [Ta09].

Um die Informationen des Messgerätes auf das Zielgerät übertragen zu können, müssen diese miteinander verbunden sein. Verbunden sind die Geräte beispielsweise über den Universal Serial Bus (USB) [Ka12a] oder sie bieten einen Datenaustausch über das Netzwerk an [Be10]. Die Art der Interprozesskommunikation zwischen Endgerät und Computer ist jedoch nicht standardisiert und häufig werden die Informationen in Hersteller-spezifischen Formaten übertragen. Zur Interpretation des Formats wird dann eine entsprechende Software auf dem Computer benötigt. Teilweise wird auch ein Software Development Kit (SDK) vom Hersteller angeboten, um eigene Programme anbinden und nutzen zu können.

2.2 Ausgabe- / Datenformate

Soweit möglich werden im Forschungsprozess Datenformate zur Speicherung durch den Wissenschaftler selbst ausgewählt. Durch die verwendeten Endgeräte oder die Software ist das Format jedoch häufig vorgegeben. Das Format ist dabei abhängig vom Anwendungsfall, des genutzten Endgeräts oder der genutzten Software. Es sollten jedoch Datenformate ausgewählt werden, die nicht nur für den aktuellen Anwendungsfall gut geeignet sind. Bei der Archivierung der Forschungsdaten ergibt sich neben der Anforderung der Sicherstellung der Integrität und Authentizität, die Anforderung diese über einen längeren Zeitraum zu erhalten. Problematisch sind beispielsweise proprietäre Datenformate, die zum Beispiel durch Endgeräte genutzt oder erzeugt werden und so nicht in jedem Fall langfristig interpretierbar sein können. Im Allgemeinen ist in der Wissenschaft schon nach den Regeln zur guten wissenschaftlichen Praxis eine Aufbewahrungsdauer von 10 Jahren vorgeschrieben [DFG98]. Die Aufbewahrungsfrist kann einzelfall- oder fachspezifisch auch weitaus länger sein. Um dies zu gewährleisten, kann das in der ersten Phase des BeLab-Projekts entwickelte Konzept, das prototypisch als Web Service umgesetzt wurde, mit einem entsprechend angebundenen Archivsystem genutzt werden. Dieses sieht eine Übergabe der Daten im Universal Object Format (UOF) vor. Darin werden die Daten, die zur beweissicheren Archivierung übergeben werden sollen, zusammengefasst [Po12].

3 Einsatz von elektronischen Signaturen

In § 371 Abs. 1 S. 2 ZPO hat der Gesetzgeber klargestellt, dass elektronische Dokumente aller Art bei der Beweiswürdigung grundsätzlich nicht als Urkunden gelten, sondern lediglich Objekte des Augenscheins sind. Dies gilt auch für elektronische Primärdaten, denn dies sind elektronische Dokumente. Besondere Beweisregeln, wie sie beim Urkundenbeweis mit Laborbüchern gelten würden, gibt es nicht. Um die Integrität der durch die Endgeräte entstandenen Daten nachweisen zu können, müssen entsprechende Verfahren, wie beispielsweise elektronische Signaturen, eingesetzt werden.

Elektronische Signaturen sind vor Gericht bezüglich des Integritäts- und Authentizitätsnachweises von besonderer Bedeutung, denn sie ermöglichen den eindeutigen Nachweis der Unverfälschtheit durch die Überprüfung von mathematischen Gesetzen. Zur Sicherstellung der Datenintegrität und Datenauthentizität von elektronischen Daten können insbesondere qualifizierte elektronische Signaturen nach dem Signaturgesetz verwendet werden. Nach § 371a ZPO gelten für elektronische Dokumente die Regeln zum Urkundenbeweis, wenn sie mit einer qualifizierten elektronischen Signatur versehen sind.

Um die Integrität von Daten mit ihrer Erhebung sicherzustellen, sind Endgeräte entworfen worden, die die Daten bereits im Gerät mit einer (qualifizierten) elektronischen Signatur versehen. Beispielsweise unterstützt die digitale Industrie-Kamera der Firma Kappa optronics GmbH [Ka12b] oder die Waage der Firma Schenck Process GmbH [Ra06] diese Funktion. Die bereits im Gerät berechnete Signatur hat den Vorteil, dass eine Manipulation nur durch den Eingriff am Endgerät selbst durchgeführt werden kann. Diese können dann beispielweise durch eine Versiegelung des Geräts sichtbar gemacht werden. Zum Beweiswert der vom Endgerät elektronisch signierter Daten siehe [Po11]. Im Forschungsprozess werden jedoch weitaus mehr Endgeräte eingesetzt, die diese Funktionalität nicht zur Verfügung stellen.

4 Datenerfassung und –sicherung

Speziell entworfene Endgeräte, wie die im vorherigen Abschnitt genannten, signieren die durch das Endgerät entstandenen Daten bereits im Messgerät und tragen so zur Sicherung der Integrität und Authentizität der Daten bereits bei der Datenerhebung bei. Dies ist die sicherste Methode und sorgt für einen hohen Beweiswert. Die Funktion wird jedoch nicht von allen Endgeräten erfüllt. Des Weiteren kann dies aus Gründen der Performance, die sich aus den technischen Details des Endgeräts oder durch Anforderungen, die sich im Forschungsprozess ergeben, nicht immer gewährleistet werden. Im Folgenden sollen technische Anforderungen und ein Lösungsansatz skizziert werden, mit dem es möglich ist die Integritätssicherung von weiteren Endgeräten zu ermöglichen.

4.1 Technische Anforderungen

Teilweise steht für Endgeräte ein entsprechendes SDK zur Verfügung, mit dem sich die automatisierte Datenübernahme und Aufbereitung realisieren lässt. Um jedoch auch für Endgeräte eine Methode zur Verfügung zu stellen, deren Hersteller kein SDK anbietet, soll ein generischer Ansatz verfolgt werden. Auch aus Gründen der unterschiedlichen Anforderungen im Forschungsdatenmanagement soll hier ein möglichst flexibles Konzept entworfen werden. Das heißt, dass die Datenübergabe vom Endgerät an das System nicht vom jeweiligen Endgerät abhängig sein soll.

Des Weiteren muss auf die Synchronisation während der Datenübergabe und der Datenverarbeitung durch das System geachtet werden. Maßgebend von elektronisch signierenden Endgeräten ist eine möglichst zeitnahe Berechnung und Sicherung der Hashwerte der Primärdaten vorzusehen, um den Zeitraum für mögliche Manipulationen an den Daten zu verringern.

Des Weiteren soll dem Wissenschaftler die Nutzung von elektronischen Signaturen zur Sicherung der Integrität und Authentizität in allen Phasen des Forschungsprozesses auf eine möglichst einfache und nicht behindernde Weise zur Verfügung gestellt werden. So können Signaturen in einigen Fällen automatisiert erfolgen. Der Umstand, dass diese Signaturen nicht durch persönliches Handeln des Signaturschlüsselinhabers erzeugt werden, schließt die beweisrechtliche Anerkennung aber nicht aus. Die Datenverarbeitung ist aber so zu gestalten, dass die automatische Signatur ein vom Signaturschlüsselinhaber initiierter und kontrollierter automatischer Prozess ist [RF04]. Nur so kann ihm auch die automatische Signatur zugerechnet werden. Auf eine manuelle Komponente kann bei der Signatur daher nie gänzlich verzichtet werden.

Daten, die in einem für die Langzeitarchivierung nicht geeigneten Format vorliegen, müssen für einen langfristigen Erhalt in geeignetere Formate migriert werden. Da im Forschungsprozess eine Vielzahl von (proprietären) Datenformaten existieren [Lu12] und diese individuelle Konvertierungsmethoden bedürfen, müssen einzelne dafür abgestimmte Module zur Formatkonvertierung entwickelt und eingebunden werden können. Dazu müssen die Datenformate identifiziert werden, um die geeignete Konvertierungsmethode auswählen zu können.

4.2 Lösungsansatz

Im Rahmen der Fortsetzung des BeLab-Projektes wurde zur Sicherung der Integrität von wissenschaftlichen Primärdaten ein System (im Folgenden Data Collector genannt) basierend auf den in vorherigen Abschnitt aufgezeigten Anforderungen entwickelt. Um der Anforderung der generischen Schnittstelle gerecht zu werden, wurde die Schnittstelle durch ein Dateikonzept realisiert und die benötigten Prozesse zur Integritätssicherung vollständig durch den Data Collector durchgeführt. Endgeräte müssen daher lediglich die erzeugten Messdaten in einem vorgegebenen Verzeichnis, das durch den Data Collector überwacht wird, ablegen. Alle benötigten Prozesse zur Integritätssicherung werden dann durch den Data Collector durchgeführt.

Um die Erweiterbarkeit zu gewährleisten, sind diese Prozesse modulbasiert umgesetzt. So können gewünschte Überwachungsmodule implementiert werden und im System über eine Konfigurationsdatei eingebunden werden. Beispielsweise wurde ein Modul zur Ordnerüberwachung implementiert, welches die Anzahl der vorhandenen Dateien im Verzeichnis überprüft. Wird eine neue Datei erkannt, ermittelt das Modul die Anzahl der im überwachten Verzeichnis befindlichen Dateien und vergleicht sie mit dem vom Benutzer angegebenen Maximalwert.

Zur Überwachung können mehrere Module gleichzeitig genutzt werden. Sie werden mit dem Start der Anwendung über die Konfigurationsdatei geladen. Im zweiten Schritt erfolgt die Anmeldung des Benutzers am System. Wurde der Nutzer erfolgreich authentifiziert, überwacht der Data Collector die vom Benutzer zuvor definierten Verzeichnisse. Der im Folgenden beschriebene Prozessablauf des Data Collectors ist in Abbildung 1 dargestellt.

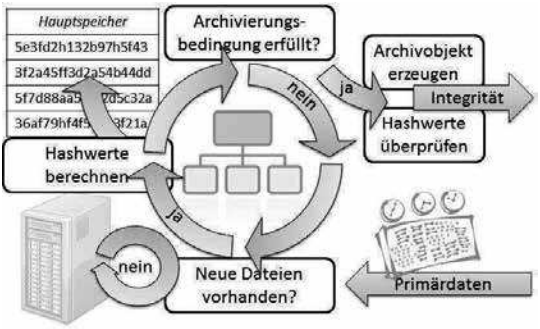

Abbildung 1: Prozessablauf des Data Collectors

Werden durch den Data Collector neu geschriebene Daten erkannt, liest dieser die Daten in den Hauptspeicher ein. Anschließend werden die Hashwerte der eingelesenen Dateien berechnet und im Hauptspeicher für einen späteren Vergleich vorgehalten. Danach erfolgt die Überprüfung der vom Benutzer angegebenen und über Module definierte Archivierungsbedingung. Für jedes Verzeichnis können mehrere Module definiert und über and- oder or-Operatoren kombiniert werden. Ist beispielsweise die festgelegte Anzahl von Dateien erreicht und ergibt die logische Verknüpfung mit anderen Modulen ein true, wird die Archivierung veranlasst. Falls die Bedingungen nicht erfüllt sind, wird das Verzeichnis weiter überwacht und die Archivierungsbedingung zyklisch überprüft. Sind alle Bedingungen erfüllt, liest der Data Collector die Daten erneut ein und berechnet die Hashwerte erneut. Sind die Hashwerte der Dateien unverändert, signiert das System die Dokumente, stellt das Archivierungsobjekt zusammen und übergibt dieses dem BeLab Web Service zur beweiswerterhaltenden Archivierung und den damit verbundenen Erhalt der Gültigkeit von elektronischen Signaturen.

Um einen Entwurf, eine Auswertung oder Veröffentlichung zu signieren, können beispielsweise die Programme, wie Microsoft Word oder Adobe Acrobat, selbst genutzt werden. Steht durch das Programm selbst keine Signaturfunktion zur Verfügung, können Programme, wie Cryptonit[1], eingesetzt werden, um Daten zu signieren. Nachteil dieser Programme ist der zusätzliche Arbeitsschritt. Mithilfe des Data Collectors können neben der Signatur von Primärdaten auch andere im Forschungsprozess entstandene Daten signiert werden.

[1] Siehe http://sourceforge.net/projects/cryptonit/.

Wie auch zur Sicherung der Integrität der Primärdaten wird dazu für die Daten des Forschers ein spezielles Verzeichnis überwacht. Speichert der Benutzer entsprechende Daten in diesem Bereich, meldet dies der Data Collector. Nach der Bestätigung des Benutzers wird die gefundene Datei signiert. Um einerseits dem Nutzer darüber zu informieren welche Datei signiert werden soll und anderseits die Signatur nicht einem beliebigen Anwender zu ermöglichen, werden zum einen die gefundenen Dateinamen angezeigt und zum anderen muss der Nutzer zur Signatur ein entsprechendes Passwort bzw. entsprechenden PIN angeben. Der Benutzer wird über die einzelnen Prozessschritte über eine grafische Nutzerschnittstelle informiert.

Die Daten, die durch den Data Collector gesammelt werden, werden vor der Archivierung in das UOF überführt, um sie dem BeLab-System übergeben zu können [Po12]. Für jedes Endgerät, das in einem dafür vorgesehenen Verzeichnis entsprechende Daten ablegt, wird ein UOF-Objekt erzeugt. Darin werden die Primärdaten aufgenommen. Die Daten sind über eine zuvor definierte Verzeichnisstruktur nach der Archivierung auffindbar. Um die Nachnutzung der Daten bzw. die generelle Nutzung der Daten zu gewährleisten, kann der Wissenschaftler das Objekt inklusive der Daten über den BeLab Web Service auslesen. So können weitere Projektdaten in das UOF-Objekt hinzugefügt werden. Die Zuordnung der Objekte zu den Projekten wird über die Projekt-ID und die allgemeine Container-ID des BeLab-Systems realisiert [Po12].

4.3 Sicherheitsaspekte

Da die von dem Data Collector überwachten Verzeichnisse von jedem Prozess gelesen und beschrieben werden können, besteht das Risiko der Manipulation der Dateien durch parallel laufende Prozesse oder den Benutzer. Diesem Risiko wird durch die Überwachung der Verzeichnisse und ein frühestmögliches Erkennen und Hashen der Dateien entgegengewirkt. Die Überwachung wird über die Verwendung des seit Java 7 zur Verfügung stehenden WatchService [Or12] realisiert. Der WatchService ist betriebssystemspezifisch implementiert und greift, falls vom Betriebssystem unterstützt, auf ein systemnahes Benachrichtigungssystem zurück. So basiert beispielsweise unter dem Betriebssystem Microsoft Windows der WatchService auf diesem System und ist so hardwarenah umgesetzt. Falls ein Betriebssystem kein solches Benachrichtigungssystem zur Verfügung stellt, wird auf den PollingWatchService des Frameworks zurückgegriffen, welcher periodisch den Ordner auf neue Dateien überprüft. Wie sich zeigte werden bei dieser Überwachungsart Dateien allerdings teilweise erst nach einigen Sekunden entdeckt, was theoretisch die Möglichkeit einer Manipulation der Datei durch Schadsoftware erhöht. Diese Möglichkeit wird bei einer nativen Implementierung erschwert, da eine nahezu sofortige Entdeckung und Sicherung der Datei durch das Berechnen eines Hashwertes stattfindet. Die Synchronisation zwischen auf einem Verzeichnis schreibenden und lesenden Prozessen übernimmt der WatchService. Ist die Archivierungsbedingung erfüllt, werden die zu archivierenden Dateien aus dem überwachten Ordner in einen temporären Ordner verschoben. Die weitere Bearbeitung übernimmt dann ein neuer Prozess des Data Collectors.

Die Hashwerte der Dateien, welche unmittelbar nach Ablegen im Verzeichnis gebildet werden, sind bis zur eigentlichen Einlagerung der Daten im Arbeitsspeicher zwischengespeichert. Hieraus ergibt sich die Möglichkeit in diesem Zeitraum die im Arbeitsspeicher abgespeicherten Hashwerte zu verändern. Eine weitere Angriffsmöglichkeit ergibt sich während der Übertragung der Daten zum BeLab Web Service. Ist keine Datenverschlüsselung vorgesehen, kann beispielsweise durch die Verwendung eines Netzwerkanalysetools der unverschlüsselte Datenstrom relativ problemlos mitgeschnitten, gelesen und sogar manipuliert werden. Daher wird für die Übertragung der Daten das HyperText Transfer Protocol Secure (HTTPS) verwendet. Als weitere Sicherheitsmaßnahme für die korrekte Übertragung zum BeLab Web Service werden die Daten bereits vor dem Versenden durch eine elektronische Signatur gesichert. Diese Signatur wird anschließend auf der Serverseite durch das BeLab-System auf Integrität überprüft. Somit können Manipulationen oder Übertragungsfehler zwischen dem Data Collector und dem BeLab Web Service zuverlässig erkannt werden.

Schwachstellen der Anwendung liegen in der eventuell nicht gesicherten Übertragung zwischen Endgerät und dem Computer auf dem der Data Collector ausgeführt wird. Des Weiteren können Manipulationen im Zeitraum von der Datenablage bis zur Archivierung erfolgen. Durch die Berechnung der Hashwerte und der erneuten Datenüberprüfung vor der Datenübergabe an das Archiv, werden Manipulationen erschwert, so dass sie nur durch ein bewusstes Handeln vorgenommen werden können. Ein versehentliches Löschen oder Verändern von Messdaten wird erkannt und dem Wissenschaftler mitgeteilt.

4.4 Rechtliche Aspekte

Die iterative Überarbeitung der Primärdaten im Forschungsprozess ist eine Herausforderung bei der Beweiswürdigung. Die lückenlose, nachvollziehbare Beweiskette lässt sich nur durch eine Versionierung der Daten erreichen. Dass bestimmte Daten in einer bestimmten Version vorlagen, kann durch ihren digitalen Fingerabdruck, den Hashwert, eindeutig nachgewiesen werden. Der mathematische nachvollziehbare Abgleich von Hashwerten ist auch im Rahmen der freien Beweiswürdigung durch das Gericht ein überaus starkes und unzweifelhaftes Beweismittel.

Vergleichbares gilt für den Einsatz digitaler und elektronischer Signaturen, die auf diesem Verfahren beruhen. Mittels fortgeschrittener und qualifizierter elektronischer Signaturen nach dem Signaturgesetz ist darüber hinaus die Zuordnung einer mit einer elektronischen Signatur versehenen Datei zu einer bestimmten natürlichen Person möglich. Durch die Möglichkeit der eindeutigen Zuordnung können Autoren- und Urheberschaft der Aufzeichnungen nachgewiesen werden. Der Beweiswert ist abhängig von der Sicherheit der eingesetzten Identifizierungs- und Authentifizierungsverfahren (etwa mittels PIN). Für qualifizierte elektronische Signaturen gelten nach § 371a ZPO sogar die für den Beweisführer besonders vorteilhaften Regeln des Urkundenbeweises.

Möglichen Angriffen gegen das System wird mit verschiedenen Sicherheitsvorkehrungen begegnet. Durch den Schutz des IT-Systems, insbesondere der Schnittstellen und Übertragungskanäle, können vom Beweisgegner nicht mehr lediglich pauschal Manipulationsmöglichkeiten gegen den Beweiswert der abgelegten elektronischen Archivdaten angeführt werden. Der Parteivortrag müsste hinreichend genug konkretisiert werden. Dies kann im Einzelfall sogar zu einer faktischen Beweislastumkehr führen, da nun sehr hohe Anforderungen an den Gegenbeweis zu stellen sind.

Fortgeschrittene und qualifizierte elektronische Signaturen sind Beweismittel für die Echtheit und Unverfälschtheit, also die Integrität und Authentizität der archivierten elektronischen Forschungsdaten. Aber auch andere Sicherungsvorkehrungen, wie die erneute Datenüberprüfung vor der Datenübergabe an das Archiv, steigern den Beweiswert. Es sind Indizien für Integrität und Authentizität der Daten. Durch diese Maßnahme kann sowohl ein versehentliches Löschen oder Verändern von Messdaten erkannt werden als auch Manipulationen durch Dritte vor der Archivierung aufgedeckt werden.

Die im Data Collector abgelegten elektronischen Daten sind mithin besonders belastbar und im Rahmen der Beweiswürdigung in der Regel als echt und unverfälscht zu bewerten.

5 Fazit und Ausblick

Die Qualität der Forschung steht u. a. im Bezug zur Qualität der Primärdaten. Durch den Nachweis der Integrität der Daten, kann diese Qualität aufrecht gehalten und nachgewiesen werden. Dazu können elektronische Signaturen eingesetzt werden und bieten, insbesondere in der Form der qualifizierten elektronischen Signatur, auch vor Gericht Vorteile. Signierende Messgeräte bieten diese Sicherheit durch einen automatisierten Prozess, werden jedoch bislang nur von wenigen Geräten unterstützt. Durch den Einsatz des Data Collectors kann die Integritätssicherung auch für weitere Endgeräte genutzt werden. Da der Signaturprozess dabei nicht durch das Endgerät gekapselt ist, sondern in einer externen Lösung durchgeführt wird, müssen mögliche Sicherheitslücken betrachtet werden. Hier wurden mögliche Schwachstellen sowohl im Programmablauf als auch in der Datenübertragung berücksichtigt.

Neben der Sicherung der Integrität der Primärdaten bietet der Data Collector dem Wissenschaftler ein effizientes Werkzeug, mit dem auch in weiteren späteren Phasen des Forschungsprozesses die Integrität gewährleistet werden kann. Durch die Anbindung des BeLab Web Services kann auch dies langfristig gewährleistet werden. Zur langfristigen Interpretierbarkeit der Daten sollen im nächsten Schritt Module implementiert werden, die eine Konvertierung von Primärdaten in proprietäre Datenformaten zu für die Langzeitarchivierung geeigneten Formaten vornehmen.

In einem im Rahmen des BeLab-Projekts durchgeführten Workshop wurde u. a. der entwickelte Data Collector vorgestellt und der Lösungsansatz diskutiert. Hervorgehoben wurde der Vorteil, dass schon durch die automatische Archivierung der Primärdaten direkt nach ihrer Erzeugung die Integrität der Daten gewährleistet wird und so ungewollte Veränderungen bei der Aufbereitung vermieden werden. Die Integrationsmöglichkeit des Systems in den Forschungsablauf durch den modularen Aufbau und die umgesetzten Schnittstellen fördert dessen Akzeptanz und wurde als eine Voraussetzung für dessen Nutzbarkeit genannt.

Literaturverzeichnis

[Be10] BestScope, BLM-280 LCD Digital Microscope, http://www.bestscope.net/_d270299724.htm, abgerufen am 08.04.2013.

[DFG98] DFG: Vorschläge zur Sicherung guter wissenschaftlicher Praxis: Empfehlungen der Kommission "Selbstkontrolle in der Wissenschaft", Wiley-VCH, Weinheim, 1998.

[Ha11] Hackel, S., Johannes, P.C., Madiesh, M., Potthoff, J., Rieger, S.: Scientific Data Lifecycle – Beweiswerterhaltung und Technologien. In: (BSI, Hrsg.): Sicher in die digitale Welt von morgen - Tagungsband zum 12. Deutschen IT-Sicherheitskongress, SecuMedia, Gau-Algesheim, 2011; S. 403 - 418.

[Ka12a] Kappa optronics GmbH, Kalypso 023C-USB, http://www.kappa.de/en/camfinder/Kalypso_023C-USB.html?p=1, abgerufen am 08.04.2013.

[Ka12b] Kappa optronics GmbH, PS 40S - 285, PS 4S-285, http://www.kappa.de/en/camfinder/PS_40S-285_PS_4S-285.html?p=1, abgerufen am 08.04.2013.

[Lu12] Ludwig, J.: Zusammenfassung und Interpretation. In (Neuroth, H., Strathmann, S., Oßwald, A:, Scheffel, R., Klump, J., Ludwig, J., Hrsg.): Langzeitarchivierung von Forschungsdaten – Eine Bestandsaufnahme, vwh, Boizenburg, 2012; S. 295 - 310.

[Or12] Oracle: The Java Tutorial – Watching a Directory for Changes, http://docs.oracle.com/javase/tutorial/essential/io/notification.html, abgerufen am 08.04.2013.

[Po11] Potthoff, J., Johannes, P.C., Madiesh, M., Rieger, S.: Elektronisch signierende Endgeräte im Forschungsprozess. In (Schartner, P., Taeger, J., Hrsg.): D-A-CH Security 2011 – Tagungsband, syssec, Klagenfurt, 2011; S. 44 - 55.

[Po12] Potthoff, J.: Beweiswerterhaltendes Datenmanagement im elektronischen Forschungsumfeld. In (Müller, P., Neumair, B., Reiser, H., Rodosek, G.D., Hrsg.): Proceedings 203 5. DFN-Forum Kommunikationstechnologien: Fachtagung Regensburg 21.-22.05.2012, Köllen, Bonn, 2012; S. 109 - 118.

[RF04] Roßnagel, A., Fischer-Dieskau, S.: Automatisiert erzeugte elektronische Signaturen, MMR 2004; S. 133 – 139.

[Ra06] Rauchschwalbe, U., Wiesmaier, A., Ludwig, C., Buchmann, J.: Digital signierte Wägeresultate, neue Wege in der Sicherung eichfähiger Messwerte. In: Wägen, Dosieren, Mischen (WDM) 2006, Ausgabe 3; S. 23 - 27.

[Ta09] Tanenbaum, A.S.: Moderne Betriebssysteme, Pearson Studium, München, 2009.