**MALIS 19.3 WPM\_T9.1: Data Science / Data Librarianship / ITPraxis**

**1. Beispiel: Revision mit Excel – keine glückliche Lösung**

- Sortieren von Metadaten

Die Revision in der Archäologischen Bibliothek erfolgt nach Möglichkeit alle zwei Jahre. Sie ist eine Methode, um festzustellen, ob alle Titel, die im OPAC eingepflegt wurden, auch im Regal vorhanden sind oder ob die im Regal vorhandenen Titel auch in den OPAC eingepflegt wurden. Dies ist für die Bibliotheksleitung wesentlich, da sie so einen Überblick über den momentanen Bestand erhält und Fehlbestände bzw. fehlerhafte OPAC-Eingaben kontrollieren und beseitigen kann. Für die Nutzer einer Bibliothek ist die Beseitigung von Fehlerquellen im OPAC essentiell, da sich die meisten Nutzer ausschließlich über den OPAC informieren, ihre Literatur suchen und auch über den OPAC den Standort des benötigten Titels in der jeweiligen Bibliothek erfahren. Wenn also ein vorhandenes Buch nicht im OPAC eingetragen ist, existiert es für den Nutzer online nicht und er kann es dementsprechend nicht verwenden. Aus diesem Grund sollten diese Titel bei der Revision auffallen und nachgetragen werden. Eine Revision wird durchgeführt, um festzustellen, ob Fehlbestände vorliegen und ob diese im OPAC auch als solche ausgewiesen werden. Dies ist der Fall, wenn Metadaten zu einem Titel im OPAC existieren, das Buch jedoch nicht im Regal steht und auch kein anderer Verbleib feststellbar ist. Fehlbestände sollten im OPAC kenntlich gemacht und ggf. nachgekauft werden, damit die oder der Titel den Nutzern wieder zur Verfügung stehen bzw. stehen oder der Nutzer darüber informiert ist, dass ein bestimmtes Buch in einer bestimmten Bibliothek nicht zur Verfügung steht.

Das Einpflegen der Metadaten in den OPAC geschieht an der Ruhr-Universität Bochum mit Hilfe der Programme Aleph und Sisis. Für die Revision wird ein Teil dieser Metadaten aus Aleph und Sisis genutzt und als Excel-Dateien von der Universitätsbibliothek zur Verfügung gestellt. In den beiden Excel-Dateien – für jeden Fachbereich eine – sind Metadaten einzelner Titel wie Autor, Titel, Jahrgang, Verlag und Signatur vorhanden. Für die Revision müssen die Daten dann aufsteigend nach der Signatur sortiert werden, weil die Signatur meinen Hilfskräften und mir mitteilt, an welcher Stelle im Regal sich das betreffende Buch befindet. Die Regale müssen wir mit dieser Liste ablaufen, damit wir das betreffende Buch in die Hand nehmen und kontrollieren können, ob es vorhanden ist oder zum Fehlbestand gezählt werden kann.

Excel

Excel hat die Vorteile, dass allgemein Daten umgeformt, kombiniert und mit ihnen gerechnet werden kann. Es ist also für Tabellenkalkulationen entwickelt worden, aber nicht, um große Datenmengen und Daten an andere Systeme weiterzugeben. Die Datenverwaltung ist in diesem Programm nur „Mittel zum Zweck“, um überhaupt mit den Daten zu rechnen und Kalkulationen vorzunehmen. Dies ist das vorrangige Ziel von Excel. Ein großer Nachteil von Excel ist die daraus resultierende lange Bearbeitungszeit bei großen Datenmengen. Dass die Datei immer nur mit dem Excel-Programm bearbeitet werden kann, und dass Excel im Laufe der Zeit bei ständig wachsenden Datenmengen an seine Grenzen stößt, sind ebenfalls nicht unbeträchtliche Nachteile. Denn in diesen beiden Dateien werden keine Daten gelöscht, sondern stetig neue Einträge geschrieben, da wir aus Forschungsgründen keine alte Literatur „entsorgen“ und stetig neue Literatur hinzukaufen.

Im Falle der Revision ist aufgefallen, dass die Daten zwar korrekt wiedergegeben werden, aber nicht optimal für mich sortiert werden können – oder ich konnte sie für mich nicht so sortieren, dass ich effizient mit ihnen arbeiten konnte. Eine Sortierung der Metadaten kann für meine Arbeitsaufgabe nicht nach Autorname erfolgen, da die Bücher im Regal nicht nach den Autoren, sondern nach alphanumerischen Signaturen aufgestellt sind. Um eine effiziente Sortierung der Signaturen zu erreichen, müsste ich alle betreffenden Daten manuell abändern, damit Signaturen in der richtigen Abfolge erscheinen.

Beispiel 1a) Ausgabe der Sortierung der Signaturen der Klassischen Archäologie in Excel:

|  |  |  |
| --- | --- | --- |
| C Baal a 1 | Harding, Gerald Lankester | Baalbek |
| C Baal a 2 | Champdor, Albert | ¬L'acropole de Baalbek |
| C Baal a 3 | Sader, Hélène ¬[Hrsg.] | Baalbek |
| C Baal b 1 | Ess, Margarete ¬van | Baalbek |
| C Baal b 2 | Ess, Margarete ¬van | Baalbek - Heliopolis |
| C Baal b 3 | Collart, Paul | ¬Le petit autel de Baalbek |
| C Baal d | Frauberger, Heinrich | ¬Die Akropolis von Baalbek |
| C Baby a 1 | Koldewey, Robert | ¬Das wiedererstehende Babylon |
| C Baby a 2 | Oates, Joan | Babylon |
| C Baby a 3 | Babelon, Ernest | Manuel d'archéologie orientale |

Beispiel 1b) Optimale Ausgabe der Sortierung der Signaturen der Klassischen Archäologie (und damit gleichzeitig, in welcher Reihenfolge die Aufstellung der Bücher im Regal erfolgt)

|  |  |  |
| --- | --- | --- |
| C Baal a 1 | Harding, Gerald Lankester | Baalbek |
| C Baal a 2 | Champdor, Albert | ¬L'acropole de Baalbek |
| C Baal a 3 | Sader, Hélène ¬[Hrsg.] | Baalbek |
| C Baby a 1 | Koldewey, Robert | ¬Das wiedererstehende Babylon |
| C Baby a 2 | Oates, Joan | Babylon |
| C Baby a 3 | Babelon, Ernest | Manuel d'archéologie orientale |
| C Baal b 1 | Ess, Margarete ¬van | Baalbek |
| C Baal b 2 | Ess, Margarete ¬van | Baalbek - Heliopolis |
| C Baal b 3 | Collart, Paul | ¬Le petit autel de Baalbek |
| C Baal d | Frauberger, Heinrich | ¬Die Akropolis von Baalbek |

Beispiel 2a) Ausgabe der Sortierung der Signaturen der Ur- und Frühgeschichte in Excel:

|  |  |  |
| --- | --- | --- |
| Fa 1 | Scollar, Irwin | Archäologie aus der Luft |
| Fa 10 | Sellnow, Irmgard | Grundprinzipien einer Periodisierung der Urgeschichte |
| Fa 100 | Herrmann, Joachim | ARCHAEOLOGISCHE DENKMALE UND UMWELTGESTALTUNG / HRSG. VON JOACHIM HERRMANN |
| Fa 111 | Gardin, Jean-Claude | ARCHAEOLOGICAL CONSTRUCTS : AN ASPECT OF THEORET. ARCHAEOLOGY |

Beispiel 2b) Optimale Ausgabe der Sortierung der Signaturen der Ur- und Frühgeschichte (und damit gleichzeitig, in welcher Reihenfolge die Aufstellung der Bücher im Regal erfolgt)

|  |  |  |
| --- | --- | --- |
| Fa 1 | Scollar, Irwin | Archäologie aus der Luft |

Fa 2 Preuss, Joachim Von der archäologischen Quelle

Fa 3 Bericht von der Restaurierung und Konservierung

Tagung

.....

Fa 10 ... ...

Fa 11 ... ...

Da bei der manuellen Abänderung der Sortierung zu viel Zeit verloren geht und es sich dabei zusätzlich um eine gehaltlose Arbeit handelt, nehme ich die manuelle Sortierung gar nicht erst vor und belasse die Liste in der Sortierung, die Excel mir anbietet. Der Nachteil dieser ineffizienten Sortierung ist ein ebenso ineffektiver Arbeitsprozess, da während der Revision die eigentliche Reihenfolge am Regal unterbrochen werden muss, um die Reihenfolge der Signaturen in der Liste einzuhalten. So müssen Wege doppelt abgelaufen werden, um die Titel zu überprüfen, die nicht in der gewünschten Reihenfolge in der Excel-Tabelle stehen. Dieser Aufwand kostet zu viel Zeit und die Gefahr von Fehlern ist groß, da Signaturen u. U. doppelt kontrolliert, dafür einige jedoch versehentlich ausgelassen werden, weil nicht mehr ersichtlich war, an welcher Stelle am Regal die Kontrolle unterbrochen wurde. So ist viel Potential verschenkt, da diese unnötigen Wege und Fehler mit der richtigen Sortierung der Signaturen vermeidbar wären.

Datenbanken

Im Hinblick auf das fehlende Potential bei Excel-Tabellen bezüglich der Aufgabenstellung „Revision“ wäre eine Datenbankanwendung eventuell die bessere Alternative, um größere Datenmengen verarbeiten zu können. Da der Vorgang der Revision alle zwei Jahre durchgeführt wird und innerhalb dieser Zeitspanne immer wieder neue Metadaten in den OPAC eingetragen werden, wächst auch die Datenmenge für die Revision stetig an. Dass bei einer Datenbankanwendung nachhaltigere Anwendungen wie das Anlegen, Speichern, Abfrage und Zugriff auf die Metadaten möglich sind, ist für das Vorhaben der Revision ein klarer Pluspunkt. Während Excel mit der Bearbeitung von großen Datenmengen schnell an die Grenzen seiner Leistungskapazität gerät, sind Datenbanken sogar darauf ausgelegt, wachsende Datenmengen aufzunehmen und eine rasche Verarbeitung zu sichern. So dauert die Speicherung und Bearbeitung von Daten nicht so lang, wie es bei Excel-Dateien oftmals der Fall ist. Wobei darauf hingewiesen werden soll, dass Excel kein schlechtes Programm ist, sondern lediglich darauf abzielt, mit den vorhandenen Daten Kalkulationen durchzuführen. Auch liegt eine Datenbankanwendung auf dem Server, der von sich aus größeres Potential zu einer schnelleren Bearbeitung mit sich bringt, weil er eine größere Rechenkapazität als der Arbeits-PC oder Laptop besitzt, auf dem das Excel-Programm gespeichert ist.

Für die Revision in der Archäologischen Bibliothek bedeutet es, dass alle bereits vorhandenen Daten wieder verwendbar sind und sich neue Daten durch einen automatischen Import problemlos in die vorhandene Struktur einpflegen lassen oder sich selbst in die Struktur eingliedern. Die Datei soll also immerfort erweiterbar sein, da die vorhandenen Metadaten nie gelöscht, sondern stetig neue hinzukommen werden. Die Anforderung der benötigten Datenbankanwendung für meine Bedürfnisse ist die effektive Sortierung der Metadaten nach der alphanumerischen Signatur in der gewünschten Reihenfolge, so dass eine effektivere Arbeitsweise als bisher möglich ist.

**Automatisierungsprozesse bei der Rechnungsbearbeitung (paperless office)**

Ist-Zustand

Am Institut für Archäologische Wissenschaften werden alle eingehenden Rechnungsbeträge im Sekretariat vorkontiert und die bearbeitete Rechnung an die Finanzbuchhaltung der Universitätsverwaltung per Hauspost gesendet. Die Rechnung muss im Original vorhanden sein, damit sie bezahlt wird. Elektronische Rechnungen, z. B. Pdf-Dokumente als Mail-Anhang zählen nicht als Originalrechnung und dürfen dementsprechend nicht bearbeitet werden. Der gesamte Prozess von der Rechnungsbearbeitung ausgehend, über die Vorkontierung im Institut und zur endgültigen Verbuchung des Betrags auf das korrekte Konto in der Universitätsverwaltung bis hin zur abschließenden Zahlung des Rechnungsbetrags an den Lieferanten kann mitunter einige Wochen dauern.

Wünschenswerter Zustand

Für die Rechnungsbearbeitung würde ich mir sowohl eine Vereinfachung als auch eine Beschleunigung durch Automatisierungsprozesse (elektronisches Datenmanagement) wünschen. Beispielsweise eine Umstellung auf ausschließlich elektronische Rechnungen, da diese schnell zugesendet, abgespeichert und wieder aufgerufen werden können. Auch die Weiterverarbeitung der Rechnung ist effizienter, da keine Bearbeitung von Papierrechnungen mit Stempel und Vorkontierungsblatt, keine Wege zum Sekretariat oder Dekanat, die die Vorkontierung vornehmen müssen, und keine Transportwege durch die Hauspost zur Verwaltung, wo die Beträge dann erst richtig verbucht werden, mehr notwendig sind. Die Rechnungsbeträge können auf diese Weise schneller verbucht werden und der Lieferant erhält nicht erst nach ein paar Wochen, sondern nach wenigen Tagen sein Geld. Somit sind sie effizienter für die Arbeitsvorgänge innerhalb der Universität, sowie zwischen Lieferant und Besteller. Außerdem sind elektronische Rechnungen nachhaltiger in Bezug auf die Umwelt, weil der Lieferant keine Papierrechnung mehr ausdrucken muss und auch der Transport der Originalrechnungen mit der Hauspost entfällt. Dies ist ein sowohl zeitlicher als auch umweltschonender Vorgang, da unser Institut nicht mehr am RUB-Campus ansässig ist, sondern vor 10 Jahren in die Bochumer Stadtmitte gezogen ist.

Vorschlag für Entwicklungspotential

Es existieren Buchhaltungsprogramme wie z. B. Collmex, die elektronische Rechnungen automatisch einlesen. Beim Buchungsvorgang zeigt das Programm das betreffende Pdf-Dokument und den zu verbuchenden Betrag an. Dann macht das Programm einen Buchungsvorschlag zur Rechnung – welcher Betrag soll auf welches Konto gebucht werden? Wenn der Anwender den Vorschlag bestätigt, wird die Buchung vollzogen und der Lieferant erhält seine Zahlung, ohne dass mehrere Wochen vergehen und die Umwelt unnötig belastet wird.