

# Entwicklung des digitalen Tanzarchivs der Pina Bausch-Stiftung

Bernhard Thull, Vera Marz

## Motivation und Ausgangspunkt

Das Werk der international anerkannten Choreografin Pina Bausch umfasst mehr als 50 Stücke, die weltweit aufgeführt worden sind. Nach ihrem Tod im Jahre 2009 hat es sich die Pina Bausch-Stiftung zur Aufgabe gemacht, dieses Werk zu bewahren. Es ist durch umfangreiches Material z.B. im Archiv der Pina Bausch Stiftung dokumentiert. Vielleicht noch wichtiger ist aber die Erinnerung von vielen Menschen in der ganzen Welt an ihr Werk. Das Werk von Pina Bausch zu bewahren bedeutet daher sowohl Material als auch Erinnerungen zu erfassen und zu erhalten.

Das Material im Archiv der Pina Bausch Stiftung umfasst Stück- und Aufführungslisten, Tonaufnahmen und ihre Transkriptionen, Regiebücher, Dokumentationen der Bühnenbilder, Fotografien und Videoaufnahmen, Kostüme, Manuskripte, Requisiten, Tänzeraufschriebe und vieles mehr. Sie enthalten Informationen und inhaltliche Beziehungen z.B. über Personen, Stücke und ihre Aufführungen, Besetzungen oder Rollen und ihre Entwicklung. Obwohl der größte Teil des Materials im Archiv der Pina Bausch Stiftung eingelagert ist, ist weiteres Material in Tanzarchiven weltweit verteilt oder nicht einmal Teil eines Tanzarchivs, wie z.B. ein Zeitungsartikel über eine ganz bestimmte Aufführung in einem Zeitungsarchiv oder ein Buch über das Werk von Pina Bausch in einer Bibliothek.

Erinnerungen sind mit Personen, Aufführungen, Szenen, oder vielleicht mit einem Bühnenbild oder mit einer bestimmten Musik verbunden und sie beziehen sich aufeinander und auf das Material in individueller Weise. Sie sind unter den Mitgliedern des Tanztheater Wuppertal Pina Bausch, aber auch unter vielen anderen Menschen weltweit verteilt. Im Vergleich zu dem gesammelten physischen Material des Archivs sind menschliche Erinnerungen von Natur aus emotional, ungenau, widersprüchlich, inkonsistent und unvollständig. Sie liefern Fragmente, die zusammen betrachtet ein Bild ergeben.

Mit dem digitalen Archiv verfolgt die Stiftung verschiedene Ziele. Das Archiv soll z.B. dem Tanztheater Wuppertal Pina Bausch dabei helfen, Stücke wieder aufzuführen, Tanzinteressierten das Werk von Pina Bausch vermitteln oder Wissenschaftlern als Werkzeug für ihre Forschung dienen.

## Anforderungen

Welche Werkzeuge unterstützen die Sammlung von Material und von Erinnerungsfragmenten und erlauben dann, daraus größere Bilder entstehen zu lassen? Die wesentlichen Anforderungen im Detail:

- Beschreibung des physischen Materials, das von Pina Bausch gesammelt worden ist. Dieses Material sollte erfasst und Beziehungen zwischen diesem Material repräsentiert werden können.
- Sammlung von Erinnerungen sowie jegliche gesprochenen oder aufgeschriebenen Anmerkungen über jede Art von Material des Archivs. Beispielsweise sollte eine Tänzerin in der Lage sein, eine bestimmte Szene eines Stücks mündlich zu kommentieren oder ein Zuschauer sollte seine Erfahrungen mit einer bestimmten Aufführung beschreiben können.
- Beliebiges Material miteinander verbinden. Beispielsweise sollte es möglich sein, eine bestimmte Requisite mit einem Manuskript zu verbinden und festzuhalten, dass die Gestaltung dieser Requisite in diesem Manuskript beschrieben worden ist. Oder ein Zuschauer hat einen seltenen Zeitungsartikel aus den siebziger Jahren gefunden und möchte ihn mit der Aufführung verbinden, von der dieser Artikel berichtet.
- Aggregation, Sortierung, Klassifikation oder eine andere Art der Verarbeitung des Materials, um es für die Entwicklung von Interpretationen, Visualisierungen oder interaktive Erfahrungen aufzubereiten. Dies könnten Untersuchungen im Rahmen von Forschungsarbeiten sein, Websites für ein bestimmtes Publikum wie z.B. Kinder, oder interaktive Installationen im Rahmen von Ausstellungen.

## Ansatz *Linked Data*

Die zu erfüllenden Anforderungen erinnern an die Situation im *World Wide Web*. Das Web ist ein verteiltes System von miteinander verbundenen Websites. Einige Teile des Web sind aufgeräumt, gut strukturiert und organisiert, andere Teile eher spontan und chaotisch. Jeder kann Inhalte beisteuern, die letztendlich ein einziges Netz von Dokumenten (*web of documents*) bilden. Das World Wide Web Consortium<sup>1</sup> (W3C) hat erkannt, dass die weiter oben genannten Anforderungen sehr verbreitet sind und für viele unterschiedliche Anwendungsbereiche gelten. Es hat daher eine neue Form der Datenmodellierung vorgeschlagen, die *Linked Data*<sup>2</sup> heißt, und die es erlaubt, das *web of documents* in ein *web of data* umzuwandeln, ohne die offene und dynamische Natur des Web aufzugeben. Die Ergänzung des *Linked Data*-Paradigmas mit Werkzeugen, die automatisch logisch konsistente Datensätze über verteilte Daten aufbauen können, erlaubt die Entwicklung des so genannten *Semantic Web*<sup>3</sup>. Es erscheint sinnvoll zu untersuchen, inwieweit sich das *Linked Data*-Paradigma als Ansatz für das digitale Archiv der Pina Bausch Stiftung eignet.

---

<sup>1</sup> <http://www.w3.org>

<sup>2</sup> <http://www.w3.org/standards/semanticweb/data>

<sup>3</sup> <http://www.w3.org/standards/semanticweb>

## Stand des Archivs

### *Experimenteller Aufbau*

Um zu untersuchen, ob der *Linked Data*-Ansatz als Basis für das digitale Archiv der Pina Bausch Stiftung geeignet ist, haben wir eine experimentelle Systemarchitektur entwickelt. Daten werden aus vielen verschiedenen Quellen gesammelt, wie z.B. aus einer vorhandenen Filemaker *Bento*-Datenbank mit Daten über Kostüme, Microsoft *Excel*-Tabellen oder einfach *Linked Data*-Dateien, die in RDF<sup>4</sup> oder Turtle<sup>5</sup> geschrieben sind und z.B. die Beschreibung der Ontologie enthalten. Der *Triple Store* selbst ist mit Hilfe des OWLIM lite<sup>6</sup> *Triple stores* in Verbindung mit der OpenRDF Sesame Workbench<sup>7</sup> implementiert. Mit Hilfe eines einfachen web-basierten Datenbrowsers, der im Rahmen des Projekts entwickelt worden ist, kann man Daten des Archivs einsehen, Daten in das Archiv einlesen und aus dem Archiv löschen. Er dient als Werkzeug, um die Korrektheit der Daten und ihrer Verlinkung mit anderen Daten zu überprüfen.

### *Prozesse und Daten im Archiv*

Da der *Linked Data*-Ansatz kein a priori-Datenmodell benötigt, um Daten zu erfassen, kann man die Prozesse der Datenmodellierung und der Datenerfassung trennen und den Entwicklungsprozess des Archivs beschleunigen. Der Datenerfassungs- und der Datenmodellierungsprozess laufen daher unabhängig voneinander. Die Eingabe von Massendaten über das Material des Archivs erfolgt zurzeit mit Hilfe von *Excel*-Tabellen. Bis jetzt wurden insgesamt fast 48.000 Entitäten erfasst und mit knapp 920.000 Tripel beschrieben, darunter:

- 466 Personen
- 54 Stücke mit insgesamt 3.081 Szenen
- 6.286 Aufführungen
- 12.951 Objekte (z.B. Programmhefte, Videos, Fotografien, Poster oder Dokumente)

Die *Excel*-Tabellen sind dabei so angelegt, dass sie keine Information über die Modellierung enthalten. Die Modellierung der Daten erfolgt erst beim Einlesen der Tabellen in die Datenbank. Dieses Vorgehen erlaubt es, das Modell jederzeit zu überarbeiten, ohne bereits erfasste Daten wieder zu verlieren.

---

<sup>4</sup> [http://www.w3.org/standards/techs/rdf#w3c\\_all](http://www.w3.org/standards/techs/rdf#w3c_all)

<sup>5</sup> <http://www.w3.org/TR/turtle/>

<sup>6</sup> <http://www.ontotext.com/owlim>

<sup>7</sup> <http://www.openrdf.org/>

## Aktuelles Modell

Die zunächst wichtige Archivperspektive haben wir mit Hilfe des Modells der *Functional Requirements of Bibliographic Records* (FRBR, [1]) realisiert. Das FRBR-Modell dient als Ausgangspunkt für die Entwicklung der *Pina Bausch Archive Ontology*, das wir nach Bedarf instanziiieren und ergänzen. Dabei verwenden wir etablierte Vokabulare, wie z.B. *Dublin Core Metadata Terms*<sup>8</sup> oder *Simple Knowledge Organization System*<sup>9</sup>, um die Einhaltung von *Linked Data*-Prinzipien sicherzustellen.

## Fazit

Wir haben genügend Daten erfasst, um die Machbarkeit und die Sinnhaftigkeit des *Linked Data*-Ansatzes für das digitale Archiv der Pina Bausch Stiftung zu zeigen. Insbesondere das FRBR-Modell konnte sein Versprechen halten und hat sich bewährt. Auch wenn die verwendete Technologie des *Triple Stores* und der dazugehörigen Werkzeuge noch relativ jung im Vergleich zu relationalen Datenbanken ist, zeigte sie sich ausreichend gereift und stabil. Dieser experimentelle Aufbau war daher ein erster wertvoller Schritt zur Entwicklung des digitalen Archivs, das durch den Einsatz eines Entwicklungswerkzeuges wie z.B. *Callimachus*<sup>10</sup> weiter professionalisiert werden kann.

Die Benutzungsschnittstelle und die Visualisierung sowohl zur Eingabe von Daten als auch zur Erkundung von Daten hat sich dagegen als eine wesentliche Herausforderung herausgestellt. Der *Linked Data*-Ansatz widersetzt sich üblichen Methoden des *User Interface Designs*, wie wir im Projekt an vielen Stellen gelernt haben. Dies hängt mit der Dynamik von *Linked Data*-Archiven zusammen, wo eine Ressource mit nur ein paar andere Ressourcen verbunden sein kann, wohingegen eine andere Ressource der gleichen Art mit Hunderten anderen Ressourcen verbunden sein kann, wie z.B. Stücke und ihre Aufführungen. Weitere Arbeiten werden einen starken Fokus auf diesen Teil des digitalen Archivs legen müssen.

Zusammenfassend lässt sich feststellen, dass der *Linked Data*-Ansatz die Anforderungen an das digitale Archiv der Pina Bausch Stiftung zu erfüllen scheint. Durch den Einsatz von *Linked Data* und die Fähigkeit des *Triple Store*, fehlende Verbindungen herzustellen, ist es möglich, lokales und verteiltes Wissen zu einer Sicht zusammenzuführen. Mit Hilfe von Vokabularen und Ontologien, die zwischen verschiedenen Vokabularen vermitteln können, ist es möglich, verschiedenen Sichten auf und Interpretationen des Materials zu entwickeln. Das Konzept so genannter Kontexte erlaubt es, Autorschaft nachzuhalten und somit Widersprüche selbst auf der Ebene von Fakten darzustellen. Die Möglichkeit, Daten ohne a priori-Modellierung zu erfassen, erlaubt es, sich dynamisch an veränderte Anforderungen

---

<sup>8</sup> <http://dublincore.org/documents/dcmi-terms/>

<sup>9</sup> [http://www.w3.org/standards/techs/skos#w3c\\_all](http://www.w3.org/standards/techs/skos#w3c_all)

<sup>10</sup> <http://callimachusproject.org>

und Randbedingungen anzupassen. Und als ein letzter, wichtiger Punkt stellt die Einhaltung der *Linked Data*-Prinzipien<sup>11</sup> sicher, dass das Archiv zukünftig leicht mit anderen *Linked Data*-Archiven verbunden werden kann.

- [1] IFLA Study Group on the Functional Requirements for Bibliographic Records:  
*Functional requirements for bibliographic records*, Final report, International Federation of Library Associations and Institutions, 2009

---

<sup>11</sup> <http://www.w3.org/DesignIssues/LinkedData.html>