

The Blue Angel in the Archives

Preserving documents by digitising them

International film and television history is documented in many places, not only in the library of the Film and Television Academy in Babelsberg (Hochschule für Film und Fernsehen HFF). But a really extensive documentation of GDR¹ film and media history exists only at the HFF. The newspaper cuttings archive on DEFA² is especially important. Here, an electronic archiving solution has been implemented to protect the unique originals from destruction by intensive research. The greatest problems arose during the preparation of the newspapers for scanning. Now research can be done electronically at the HFF in these records.

Babelsberg Film City

It is hard to imagine a place with more historical associations: The Blue Angel, the Nibelungs, Metropolis or Lubitsch, Murnau, Lang or Lilian Harvey, Marlene Dietrich, Zarah Leander. The name of Babelsberg evokes memories, dreams, and a flood of pictures for any cinema enthusiast. Film production began here in a glass studio in 1912, the year the Titanic sank. Since then, Babelsberg has stood for national and international cinema. After the legendary Ufa³ era and the DEFA years during the time of the GDR, the French concern Compagnie Générale des Eaux (known today as Vivendi) has started the third chapter in the history of Babelsberg. Today, television production shares this home with filmmaking, and about 3000 permanent and freelance staff bring the cinema city to life every day.

Film and Television Academy

The Film and Television Academy "Konrad Wolf" (HFF) in Potsdam-Babelsberg will soon be integrated in the cinema city, and it can look back on an equally turbulent history. The HFF is a state run academy near Potsdam in the federal state of Brandenburg and the oldest German film academy. As a result of the artistic aptitude tests, only 80 to 100 students out of 1000 applicants are admitted each year. The education is interdisciplinary in ten areas of study and enables the graduates to master the complex practical challenges of film and television.

As in other academies, access to the stored knowledge of an academic library plays an important role. The HFF library exists since 1956 and is organised in the four divisions of book department / general knowledge, newspaper cuttings archive, AV media archive and videos. In 1993 the Central Film Library in Berlin and the Babelsberg academy library were amalgamated. Now the entire film and media literature published in the GDR is in the HFF

¹ GDR - the German Democratic Republic - the Eastern part of Germany before reunification.

² DEFA - the former East German state film studios (Deutsche Film Aktiengesellschaft).

³ Ufa - (Universum Film Aktiengesellschaft) German film studio, state owned from its foundation in 1918 until 1921, private until 1937, then state owned again until the facilities went to DEFA in 1949 after the partitioning of Germany.

library. Furthermore, it has the almost complete collection of DEFA film scripts, literature from Eastern Europe in the original languages and cinematic literature prior to 1945. There are also subscriptions to 165 current periodicals on related subjects.

Newspaper Cuttings Archive

The HFF library's newspaper cuttings archive is something special. The collection has grown and been made more accessible since 1996, and contains 1.6 million articles from the German speaking daily papers about film, television and new media. The GDR is represented equally with the West German and international press. The storage profile encompasses articles catalogued by country on national developments in film, entertainment and documentary films (including critiques), television films, TV series and TV journalism, national and international festivals. It contains biographical articles with 5 500 entries and specialist material about the HFF. But it does not stop there. The archive contains a subject related collection with articles in the areas of film and television science, media history, media and communications research, production, lending, technology, cinema technology, international radio and television companies, and media training centres. Particularly important are the special collections of DEFA film and Eastern European film. This collection alone, which contains about 100 000 articles, makes the archive unique. It is the only archive world-wide which possesses such a complete collection of material on all DEFA entertainment films since 1946 and many DEFA documentaries, about the history of the DEFA studios and many facts about the DEFA filmmakers. This makes it a valuable and authentic source for national and international media research. The newspaper documents were previously kept in folders sorted by subject, and these in turn ordered in superordinate envelopes. The documents have different shapes and formats, depending on how the article was printed in the newspaper. In cases where an article was not presented contiguously due to paging or other printing reasons, the parts were put together and glued to a suitable substrate.

New Entries in the Electronic Archive

Up to now, research was done in the originals of the historical archives. Before 1996, this was the only way. The facilities of the archive were mainly used by students and teachers of the HFF. But students from other institutes, and national and international film and media scientists, as well as other interested enquirers researched frequently in the newspaper cuttings archive. To better fulfil the increasing numbers of enquiries, the HFF decided to build an electronic archiving system to improve the research facilities. After a one-year preparation time and a national request for tender, the archiving system EASY-Archive was purchased from Easy GmbH of Mühlheim. A system with two workstations was installed in 1996. Besides the scanning station there is a research station which is used not only by the library staff for external research requests, but also directly by users themselves. Access to the archives from outside, e.g. on the internet, was not foreseen, as copyright laws required research to be done exclusively in the rooms of the archive.

The system is based on a full text database and new clippings are captured digitally from a scanner, converted to ASCII text with an OCR program and then included in the database.

The indexing is done manually using a thesaurus. In addition to the text input, the scanned images are stored in TIFF files. All new entries are stored on WORM discs accessed by a jukebox. About 12 000 new newspaper articles a year are archived in this way. A retrospective entry of the existing stocks could not be done at that time due to cost and capacity restraints.

Scan Solution for the DEFA collection

Due to intensive use, the physical condition of this unique specialised collection has deteriorated so much in recent years, that a digital archiving solution had to be considered to preserve the originals. Of course this was also to be integrated in the existing database. However the preparation of the originals for scanning and the scan itself turned out to be very difficult, because the newspaper cuttings, some over 50 years old, presented many different problems. The HFF decided to do the preparation themselves and leave the actual scanning to professional suppliers. The contract was won by the Microfilm Center Klein (MIK) from Berlin. MIK has been active in this branch for over 20 years and is one of the largest providers of microfilm and document management services, with a staff of over 100.

Extensive tests had to be carried out with generally available scanners for the condition, format, size and other characteristics of the newspaper clippings. However the results were not satisfactory. The best results were obtained with the new scan system "Bookeye" from the company ImageWare Components GmbH in Bonn. Microfilm Center Klein therefore installed a Bookeye scanner from ImageWare. Bookeye uses an innovative scanning concept with integrated scan and print electronics, with which the originals are digitised by an overhead scanner without physical contact. The system thus makes it possible to scan (and copy) directly from paper files without removing the pages, and also from books, periodicals and newspapers up to a format of DIN A2. Even three-dimensional objects such as architectural models can now easily be copied. The facility to scan up to DIN A2 was very important for the newspaper cuttings.

A Lot of Work

50 000 newspaper cuttings from the existing DEFA collection had to be entered into the electronic archive. The correct preparation of the documents for scanning was the problem needing the most work. The main problems were: variable, often very poor paper and print quality, stampings that are difficult to scan, red ink that cannot be read by scanners, perforated date stamps that cannot be seen by the scanner, articles in white or red ink on a black background, titles in red ink, articles in large formats, incomplete articles that had to be put together later, finding suitable user-friendly subject and article indexing categories. The HFF did extensive and very time consuming preparation work to ensure an optimum scan result. For example stamps were retraced, dates and partially destroyed articles were restored, source references were researched, red and other coloured titles and texts were rewritten, irrelevant photographs removed, redundant articles rejected, indexing criteria for the individual categories created and continually checked, associations made, title lists checked and extended. But that was not all.

Texts with a dark background were copied to obtain a higher contrast for better scans. This particularly applied to press magazines that often had to be completely recopied before scanning.

As the stocks contained very old articles from the nineteen-forties that had been severely worn by years of intensive use, many articles had to be glued together again. Incomplete articles often came to light, but these could mostly be reconstituted with copies of documents from the city and state libraries. However some articles had to remain incomplete and were identified as such in the archive. Not all damage to articles stemmed from wear and tear during research. The filming of a large part of the stocks several years before caused substantial damage to many articles, which was tedious to repair. Some newspaper articles had been completely mutilated and forced into a certain format, or glued together in such a way that they were no longer legible. Often the readability was very much reduced by these reorderings. In these cases the complete articles were reconstituted in their original form by searching them out of bound copies of the papers.

All in all, this effort resulted not only in a high quality of the digitised results, but also in the user friendliness of this important part of the archives, not to mention the long term salvation of the documents.

The newspaper documents of the DEFA collection were integrated in the existing electronic archive as images. They were tagged with search indices so that it is now possible to research on criteria like film titles, people and other cinematic data, which are contained in the catalogue. For historical reasons, the original documents are all stored in acid-free envelopes in a separate archive.

The digitising of these newspaper cuttings, some of which are over 50 years old, has guaranteed the HFF not only the preservation of the irreplaceable originals but also the possibility of fast and unrestricted research in the stocks. It has also demonstrated that very careful preparation and specialist technology (in this case the ImageWare scanner technology) are needed to implement such solutions successfully.