

Eva Kraus & Regina Retter

How to Digitise 15 Million Press Photos: The STERN Photo Archive at the Bayerische Staatsbibliothek

Abstract

In 2019, the Bayerische Staatsbibliothek (Bavarian State Library) took over the analogue archive of the weekly magazine *stern*. Digitising its 15 million pictures poses challenges in terms of logistics, process management and indexing that go beyond traditional librarian methods. In order to make the archive digitally available to the public, the Bayerische Staatsbibliothek is exploring new ways of organising digitisation workflows and indexing. The article describes the challenges inherent to the special nature of the archive and points out ways in which the library is currently addressing them.

Keywords

Stern Photo Archive, mass digitisation, photo journalism, digitisation workflow, indexing

From Journalistic Source to Historical Material

For a long time, the weekly magazine *stern* was *the* magazine for photojournalism. Founded in 1948, it quickly became one of the world's most influential magazines in the 1960s and 1970s, developing a unique visual power through its photographs.¹ Its photo archive shaped the visual memory of Germany and is one of the most important collections of photojournalism in the world.

Around 2001, *stern* turned from analogue to digital photography, like many other magazines worldwide. In the following years, *stern*'s journalists used the analogue photo archive less and less. As a result, it was partially transferred from the *stern* offices to an external storage facility in Hamburg in 2015-16. Around that time, researchers and external users began to call for easier access to the *stern* pictures and for their digital preservation within an experienced archival institution. The archive thus transformed from journalistic resource to a collection of historical interest.

The Bayerische Staatsbibliothek (Bavarian State Library) recognised the cultural-historical significance of the archive and the opportunities it offered. After consultation with the publishing house Gruner+Jahr, owner of the STERN photo archive, the library successfully received the entire collection as a donation. In 2019, eight trucks drove from Hamburg to Munich, transporting around 15 million negatives, slides, and prints. On accepting the donation, the Bayerische Staatsbibliothek promised to make the STERN Photo Archive digitally available to the public and to index the pictures. Since then, this task has become one of the most challenging projects within the library.

The digitisation and indexing of the *stern* archive is unique in several ways: First and foremost, the photo archive of *stern* stands out due to the sheer number of pictures. Secondly, the STERN Photo Archive covers a wide time span: the entire second half of the 20th century, to be precise, the period between 1948 and 2001. Thirdly, it includes a huge range of topics: domestic and foreign politics, culture, celebrities, cars, crime, catastrophes, socio-political questions, and many other issues. Fourthly, the archive contains the work of more than 100 photographers, both staff and freelance, some of whom have won prestigious awards for *stern* photos made around the world. Finally, the importance of the *stern* archive for the history of photography cannot be overemphasised.

Goals

What does the library hope to achieve with this project? Overall, we hope to make the *stern* photos digitally available for both commercial and academic use. This means that users can licence individual pictures for publications, exhibition, and other purposes and that researchers can access the digitised material.

At the start of the project, Klaus Ceynowa, Director General of the Bayerische Staatsbibliothek, noted that the library would treat the STERN Photo Archive as ‘pure content’. Its analogue character was seen as simply a ‘temporary state of materiality.’² Through digitisation, we can, all in one go, preserve this content for future generations, make it accessible throughout the world, and enable new forms of research. We will not limit ourselves to the highlights of the collection, the iconic photos from all over the world that made *stern* famous, but we will present the archive in its entirety.

The wealth of material, put together over several decades, will enable research questions that move beyond the development of individual photographers, in terms of interest and artistic style. For example,

researchers could use the archive to study the working methods and daily practices of photojournalism in general. On a meta level, researchers could use the material to observe changes in aesthetic preferences or shifts in the choice of topics. Making this content digitally available will allow researchers to examine such questions through new methods in the digital humanities, such as data mining or deep learning. However, we still have a way to go before we reach that point.

Where to Start? Initial Challenges

As previously mentioned, we face one major challenge in making this photo archive digitally available: its vast quantity. We identified three main obstacles that had to be overcome before we could start the digitisation process.

The main difficulty lies in the heterogeneity of STERN Photo Archive. The photos differ widely in their materiality, their cataloguing status, and sometimes even their state of preservation. The archive consists of three parts. The negative archive is the largest part and contains ring binders with approximately 12 million negatives and contact sheets. The second, much smaller, part contains prints and slides (up to three million), which was divided in two parts subsequently, when the archive moved to an external storage facility. The 'A archive' contains prints and slides that were considered useful for future use, due to their subjects, whereas the 'B archive' contains photographs of less recurrent topics. Within this basic structure, the photographs in the archive are subdivided in published or non-published photographs, made by staff or freelance photographers, and in colour or black and white photographs, the majority being the latter.

The material was not catalogued in a uniform way. Between 1948 and 2001, archiving practices constantly evolved, leading to discontinuities in archival systematics. Moreover, some of these changes were only partially and retroactively applied to certain sections of the archive. Consequently, the depth of available descriptive information (metadata) varies widely, making research in the archive a highly complex task. It requires navigating through two databases, a card catalogue, and several lists and tables. Additionally, the available documentation is not necessarily complete or correct. Some photographers retrieved some of their pictures from the Gruner+Jahr archive, and either stored them privately, or donated them to other archival institutions. In close coordination with photographers and their heirs, the Bayerische Staatsbibliothek successfully reassembled most of the missing parts of the archive. However, there are still gaps to be filled. Reintegrating material that

was previously returned to photographers can be a tricky process as the photographers frequently rearranged the pictures in ways that are not necessarily consistent with the rest of the archive.

As with any digitisation project of relatively recent material, legal issues are a major concern. When the Bayerische Staatsbibliothek acquired the collection, the legal status of most of its content was unclear. The original contracts between Gruner+Jahr and the photographers were either lost or never existed in written form. In the rare cases where contracts did exist, they were ambiguous and open to interpretation.

Because German copyright law applies to every picture until 70 years after the death of the creator (photographer) we had two options. We could either wait several decades before putting the digitised photographs online or try to secure usage rights for all pictures in possession of the Bayerische Staatsbibliothek. This last option entailed negotiating agreements with approximately 80-100 different photographers and their heirs.

Finally, we were confronted with a lack of time and of available personnel. The transfer of the STERN Photo Archive to the Bayerische Staatsbibliothek attracted public attention and led to high expectations about the digital accessibility of the pictures. These expectations have to be met by a team of only five librarians who oversee the project.

Our Solutions

What solutions did we find to resolve these challenges? Above all, it quickly became clear that we needed to narrow down the material to a manageable number of pictures before we could start the digitisation process. Producing a homogeneous subset of pictures would allow us to digitise them and make them available within a reasonable time frame. We decided that pictures in this subset should have the same material form, should be described by metadata, and should not have copyright concerns.

The ideal starting point for this subset proved to be the negative archive, in particular the material that was produced after 1973. In that year, the STERN Photo Archive began to print detailed information about the scenes depicted on the back of the contact sheets. As a result, we can link the images to a specific photo essay and automatically extract several metadata points. Every contact sheet already had a systematic call number from Gruner+Jahr, making it individually identifiable and traceable during the digitisation process. This is not always available with older material. Therefore,

the corpus of the initial digitisation project was set at 35 mm negatives that included contact sheets annotated on the back.³

As described above, the usage rights of the photographs in the subset needed to be secured. The cultural, political, and historical importance and relevance of the material, as well as the high demand for reproductions, called for immediate action instead of waiting for copyright protection to expire. Therefore, the Bayerische Staatsbibliothek negotiated standard contracts with long-term employed *stern* photographers or their heirs. These agreements permit the library to digitise all photographs, to display them in a digital form on the internet, and offer them for use. Photographers will receive a 50% share of the proceeds. Since October 2021, 30 photographers have signed an agreement on the use of their photographs in the wider context of the *stern* archive. Fourteen of them are represented in the initial digitization project. Their photographs make up more than 60% of the entire archive.

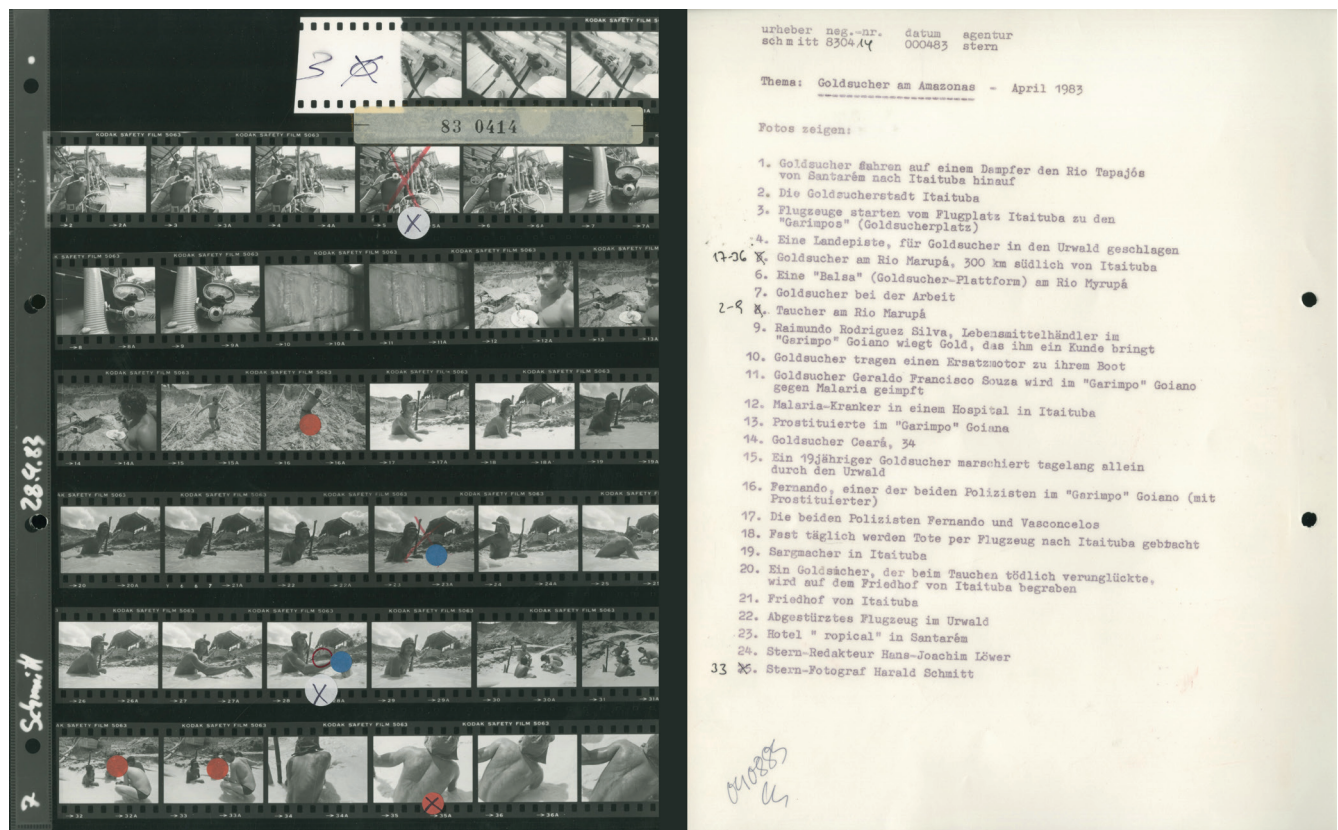


Figure 1. Front and back of a contact sheet. Excerpt from the photo report "Gold Prospectors at the Amazon," photographed by Harald Schmitt, April 1983. © Bayerische Staatsbibliothek/STERN-Fotoarchiv/Harald Schmitt.

Even though the first phase of the digitisation project involves ‘only’ three million images, the core team of the STERN Photo Archive had to seek help from external service providers. After a Europe-wide tender, three companies were selected to process the negatives by the end of 2025. These companies will produce digital copies, label them with specified file names, deliver them in JPG and TIFF formats, and transfer the negatives to new archive sleeves before returning them to the Bayerische Staatsbibliothek.

The contact sheets will be scanned in-house by a team of student workers (eight people at the time of writing). They started in February 2021 and will digitise around 15,000 contact sheets over the next four years. They also prepare the negatives for transport to the scan providers, perform quality controls, and record formal metadata.

The Digitisation Workflow

After we determined the extent of the initial digitisation project and found the people and companies to support us, we felt well positioned to start the process. To ensure the success of the project, it is crucial to closely monitor the workflows and outcomes, especially given the extent of the project and the number of physical artefacts (photographs) involved.

The negative archive has a specific structure. Negatives and contact sheets are stored together in ring binders, each marked with a STERN Photo Archive signature. Each contact sheet or corresponding negative sleeve has a contact sheet number. Individual images can then be assigned an image number based on the contact sheet number. In addition, the photo essay to which the images once belonged serves as a content-related container, which might hold several contact sheets. This last feature is only significant with regard to the indexing of the images, not to the digitization workflow. Finally, the film to which individual pictures belong can be deduced from the image number. The ring binder is the critical unit for this process. Every binder goes through five phases, which we will all closely monitor.

Firstly, each ring binder is prepared for digitisation. A student worker evaluates its content, notes the contact sheet call numbers, counts all negatives, and marks possible damage to the material. This crucial step forms the basis for the digitisation process of the scan providers. Additionally, it helps us to manage later stages in the project as it provides insight into the composition of the collection. We need this overview to ensure that the stream of material sent

to the scan providers remains stable over time, despite the fact that each binder contains a different number of pictures. During the evaluation process, the student workers also separate negatives from contact sheets. The negatives are stored in acid free archival boxes and sent to the scan providers, while the contact sheets remain at the archive in their original ring binders.

During the second phase, the in-house team digitises the contact sheets, both front and back, the text on the back page being retrieved via OCR (Optical character recognition). Supplementary material, such as registers, lists of borrowers, or slides, is not digitised and remains in the folder.

In the third phase, we monitor the quality of digital scans delivered by the scan providers and prepare them to be uploaded into the picture database or in long-term digital storage. Monitoring includes automated checks, such as generating hash files to avoid data loss and validating the TIFF files, as well as manual checks of file names and scan quality.

During the fourth phase, basic metadata are added to the digital scans. At this stage, we only add a minimum of formal information to the scans, such as the photographer's name, colour, call number, and whether the picture was highlighted on the contact sheet (more on this later). The most delicate decision that we need to take at this point is whether to release the scanned picture for online publication. While we digitise all pictures, photographs that might infringe third party rights

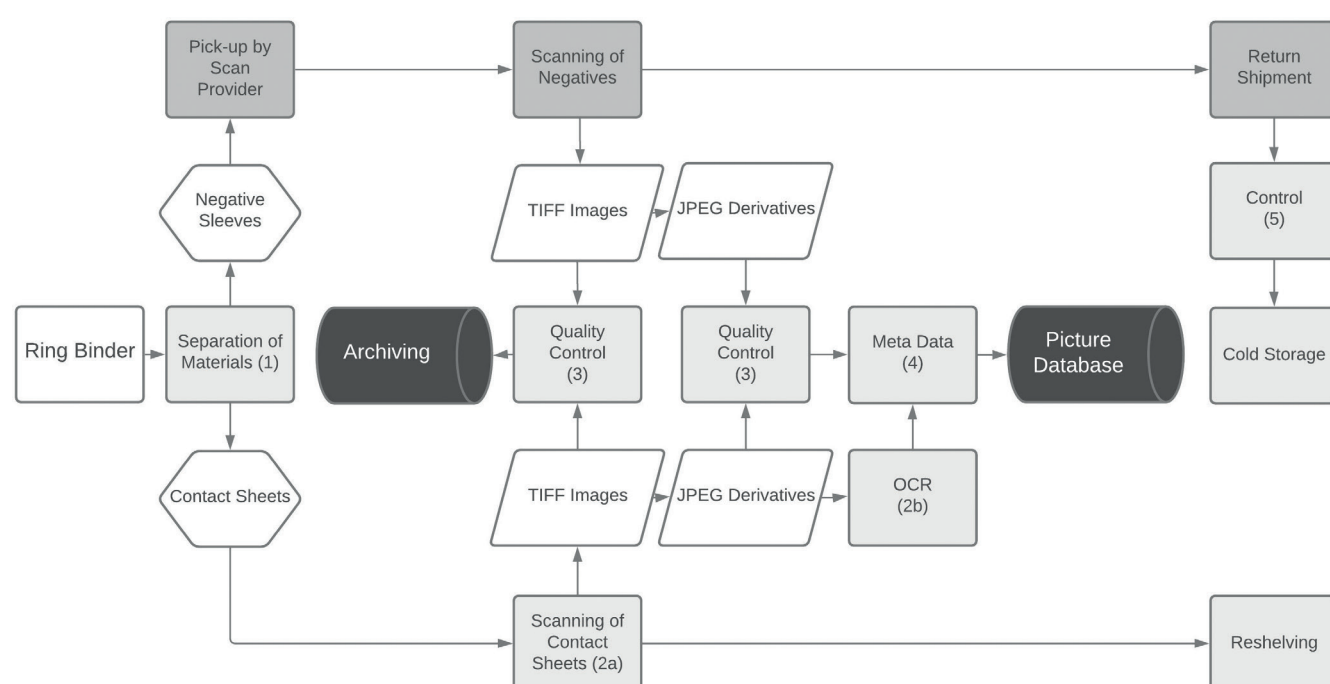


Figure 2. Digitisation Workflow. © Bayerische Staatsbibliothek.

need to be pixelated in the database. This is true for both the individual picture and its reproduction on the contact sheet. These legal restrictions concern not only copyrights but also privacy and portrait rights. Considering the historical, political, and social relevance of *stern*, it seems appropriate to grant access to as much content as possible. We make exceptions for pictures of children, pictures taken without consent, or pictures depicting people in a possibly demeaning situation or context. In the end, the decision to make an image available revolves around the question whether the photograph is historically significant, and not purely sensational, and whether it is documentary in nature, and does not cater to voyeuristic impulses. We have developed detailed guidelines to help us make these complex distinctions.

In the final phase, the student workers check the returned material for its integrity and reintegrate it into the archive. The boxes with negatives are sent to cold storage, and the old, now empty negative sleeves are put back in the ring binders with the contact sheets. This is important, as the ink on the backside of the contact sheets has transferred to the sleeves in some folders, making them a valuable resource for reconstructing the original backsides.

Database and Indexing

The digitisation of the photographs is only the first step towards digital availability. The indexing and digital presentation of the images is equally important. To this end, a new database is being set up for the STERN Photo Archive. It consists of two parts: an internal database (backend) to index the digitised pictures, and an external customer portal (frontend), where users can explore our inventory and place order requests. The data accessed by both systems - both the digitised material and the metadata - is stored on the servers of the Leibniz Supercomputing Center in Munich, which hosts all digital collections of the Bayerische Staatsbibliothek.

The vast number of pictures requires a special procedure for indexing the digitised *stern* photos in the database. The traditional indexing process for photos in the image archive of the Bayerische Staatsbibliothek provides a detailed description of each picture, including location, occasion, and objects shown, as well as the names, posture and clothing of those depicted. Such in-depth indexing is impossible due to the enormous size of the STERN Photo Archive.

In contrast, the indexing of the STERN Photo Archive does not start at the level of the individual photograph but of the photo essay, which combines related images by the same photographer in terms of content and time frame. In the original archive at Gruner+Jahr contact sheets and related images were already sorted and indexed according to the photo essay. Several metadata points, such as the photographer, time, place, and subject of the photo essay, can be found on the backs of the contact sheets. After these data points have been identified at the level of the photo essay, they are automatically transferred to all the related individual pictures. Due to time constraints, we cannot provide more detailed information on the photo essays or their individual pictures. However, we will improve the original cataloguing level to library standards by using the Integrated Authority File (GND) managed by the German National Library. A search within the GND, integrated into the database, makes it possible to index geographies, persons, institutions, subject terms, and events in a standardised way. Automatic updates guarantee that the GND records used are up-to-date. Using GND vocabulary also has a visual advantage for the presentation of the pictures in the customer portal. Because geographic data within the GND are usually provided with geographic coordinates, the photo essays can be located and displayed on a map of the world or individual countries.

As far as we know, another rather unique feature of our customer portal will be the integration of the digitised contact sheets, which will allow users to contextualise the individual image in its historical setting. Each contact sheet shows the sequence in which the pictures were taken, providing information about the actual process of taking the photographs. The contact sheets also frequently include coloured dots or other forms of highlighting placed on certain pictures by either the photographer or the *stern* editorial staff. Descriptions of the subjects depicted can be found on the reverse, in varying detail.

Typically, all contact sheets belonging to a specific photo essay contain an identical typewritten text on their reverse. Additional handwritten annotations may further specify the content of the individual contact sheet. For data protection reasons, some reverse sides will appear partially blurred in the database. Making the contact sheets available online will help users to identify the content of individual photos and allow them to track editorial decisions that are documented on the sheets. By making the individual pictures *and* the corresponding contact sheets available online, we hope to facilitate new research approaches.

Perspectives

The digitisation schedule is tight. The student team has been working since February 2021; the first service provider started negative digitisation in September 2021 and will be followed by a second and third service provider in March and May 2022. At full capacity, up to 17,000 images can be processed per week and around 65 photo essays catalogued. The launch of the customer portal is planned for late 2022. It will initially start with a six-figure number of pictures and gradually expand to three million images over the next four years.

This first step is only the beginning of the project ‘digitization of the STERN Photo Archive’. Three million slides and prints, as well as nine million negatives from the early *stern* years or from photographers with whom we do not have any usage agreements yet, are still waiting to be digitised. The experiences from this first sub-project will surely help us to develop methods and tools to address these new challenges.

Notes

1. Stefanie Nannen, *Henri Nannen: Ein Stern und sein Kosmos* (München: Bertelsmann Verlag, 2013), 15, 19.
2. Klaus Ceynowa, “15 Millionen Bilder für die ‘Ewigkeit’. Das stern-Fotoarchiv an der Bayerischen Staatsbibliothek,” *Zeitschrift für Bibliothekswissenschaft* 6 (2019): 271-279, 273.
3. True homogeneity, however, is unfortunately impossible. We learned that even within this restricted corpus, the material is full of surprises. We constantly stumble upon new exceptions, for which we need to establish rules, because, for example, films lack numbers, negatives have been cut out or rearranged, contact sheets have been cut into pieces and stapled together or are not equipped with the information they usually carry, etc.

Biographies

Eva Kraus is a research librarian and member of staff at the Department of Maps and Images of the Bayerische Staatsbibliothek in Munich. She has an academic background in history, with a specialisation in 20th-century German history. At the STERN Photo Archive, she is mainly responsible for the database and indexing processes.

Regina Retter is a research librarian and Deputy Head of the Department of Maps and Images of the Bayerische Staatsbibliothek in Munich. A political scientist by training, she is fascinated by the STERN Photo Archive and its rich material concerning the political imagery of the Federal Republic of Germany. She is responsible for the digitisation workflow of the STERN Photo Archive project.

TMG Journal for Media History

Volume 25 No (1/2022)

DOI

<https://dx.doi.org/10.18146/tmg.830>

PUBLISHER

Netherlands Institute for Sound & Vision

COPYRIGHT

Each article is copyrighted © by its author(s) and is published under license from the author(s). When a paper is accepted for publication, authors will be requested to agree with the Creative Commons Attribution 4.0 International License.