Digitization of the
Samuel A. Goudsmit papers
Melanie J. Mueller
Archivist
mmueller@aip.org
301-209-3179
Samuel A. Goudsmit Papers

38 linear feet of manuscript materials – correspondence, reports, lectures, notebooks
Our most popular and most used collection in the archives
Showing wear and tear, and in need of some attention
Document physics research from the 1920s through the mid-1970s.

Who was Samuel Goudsmit?
Physicist
Co-discoverer of electron spin, with George Uhlenbeck, in 1925
Helped establish the first programs in theoretical and quantum physics in the U.S.
Chief of Scientific Intelligence, Alsos Mission – followed Allied troops into Germany after WWII to discover the status of German atomic bomb research
Editor-in-chief of the American Physical Society; editor of the Physical Review
We chose the Goudsmit papers because it is our most popular and most used collection. Because of all the use, it was beginning to show some wear and tear. In the past, we might have considered the papers as a candidate for microfilm but we wanted to explore digital options. While the grant we received was for access NOT preservation, the side benefit is that the digital surrogate will cut down on handling and wear of the original collection.

We did NOT use any selection criteria – we were trying to follow the mandates of mass digitization as best we could. We did skip over some obvious duplication – we identified and marked those items to skip during the document preparation – as well as published or widely available items. We did not assign any additional metadata, aside from a single unique filename for each image. The filename consists of a string of numbers that reflect the hierarchy of the finding aid and allow us to keep the images in the exact same order as the physical papers, without adding description or other metadata to each image.

The project was partially funded by a grant from the NHPRC – it was a grant to improve access to collections. The grant money funded the costs to the scanning vendor, and we matched those funds to cover the staff and other costs accrued.

We did NOT perform the document scanning in-house. During the grant application process we researched many local and regional firms. We narrowed the field to three candidates based on cost and recommendations from other archives. We chose one local
scanner who met our qualifications.

We did not seek additional permissions from the thousands of individual correspondents in the collection. We are relying on interpretations of fair-use to fulfill copyright needs. We will, of course, remove any correspondence from the online collection upon the request of heirs of the correspondents. In the interest of providing increased access to the collection we did not have the staff and money to seek individual permissions. There are links to our access application and rules and regulations governing citing and quoting and copying from collections. We feel safe in assuming that researchers who use this collection will obey the same rules and researchers who visit the collection in person.
Well, things didn’t quite go as smoothly as we planned...

Workflow: how we thought it would happen...

• Hire project assistant
• Prep documents for scanning
• Receive scans from vendor in batches
• Begin with full quality review of scans, then progress to spot-checking
• Work with AIP web design team to create website
• Send to test audience
• Go live with online collection
• Run victory lap around the archives!
We hit a snag immediately, as the vendor that we contracted with during the application process had been bought by a much larger company by the time we got our funding. We had to spend months renegotiating our contract back to the original terms. We also were hit with an internal hiring freeze which delayed hiring the part-time project assistant.
After the first delivery of scanned images, we immediately began noticing problems. Poor quality scans, upside-down scans, mis-ordered original folders and general mayhem – problems that, we would come to find out – would plague us until the final weeks of scanning. We had near-constant communication with the vendor, weekly (or more) meetings at the vendor location and countless re-scans and still we had errors.
Despite dealing with the errors, we had some good progress in-house with the design of the online interface for the collection. By the end of 2010 we had a finished website to deliver the collection, but no images to test out the process.
Mid-way through 2010 we finally reached the point where we were getting consistently good results from the scanning vendor so we had them RESCAN THE ENTIRE COLLECTION. All of it. The amount of errors in the previous batches were so great that it was going to be faster and look more consistent to simply scan the entire thing over again instead of trying to correctly replace the bad scans.
After a few no-cost extensions to the deadline, we were finally done with the scanning. The website went live in April, we sent the link out to a volunteer test audience of about 45 historians and scholars and library and archives professionals, then we turned our final report in to NHPRC.
Fun facts about the project

• 10 months of research and prep before even applying for the grant

• 28 months of work once the grant started

• 1,800 hours of document prep and image checking (that's 225 8-hour work days!)

• 69,000 digital images in the collection

• 17 GB (master TIFF, JPGs and JPG thumbnail images)

• 3000+ metal fasteners removed during document prep...
I still have all this trash sitting in paper coffee cups in my office. It took so much time to prep the documents that I can’t bring myself to throw the garbage away...
Lessons learned

• Vendor selection
• Review by historians
• Project assistant
• Publicity

We spent an excruciating amount of time choosing the vendor to perform the scanning. The company came with glowing recommendations. And still, life can intervene in ways you would never imagine. The company we chose was purchased by a much larger company, one that was no longer focused on archival quality scanning and focused more on mass scanning of paystubs and legal documents where the final image does not need to be perfect. It took many months, nay, a year, of near constant badgering before we got the results we needed.

We sent a sample of the website to a small group of volunteers early in the design process, to make sure we were on the right track with the initial design. Once we had the full site up and functional, we sent the website out to a larger more diverse test audience. We ended up receiving some invaluable feedback on the design, functionality and navigation of the website. We were able to incorporate some changes immediately, but the rest will be added during an overall redesign of all of our websites in the coming year.

The part-time project assistant proved to be the glue that held this project together. She worked 20 hours per week and did NOTHING but proof and compare scanned images and original documents. Since portions of the collection were scan 3+ times before they were correct, this meant she had to re-do months of work in the end. This was a very tedious job and we were explicit when interviewing candidates that this was going to be no easy task. Finding a person who was diligent and exacting and focused, even in the face of extreme tedium, was the key to our success.

We are lucky to be part of a larger organization which had a department dedicated to media and press. We had identified publicity sources specific to archives and historians (listservs, newsletters, announcements, etc.) but our marketing team was able to draft a press release and push it out to newswire organization to get picked up by news all over the world. We will be monitoring the web traffic to the site, to see which of these sources
were the most effective.
Dear Esther,

Are you glad to be back in school again? Did the new children in your grade and one of you leave one thing?
I had dinner at a home with two little boys, of six and nine, and I showed them your picture. How they want to live in America and visit you!

Please write me soon, my address is:
S. Simmons
305 S. 8th St.
San Antonio, Texas
A lot of kisses for you and from Mummy.

From your:

Daddy

Greetings for Mary.
Alsos mission
Jail 'Mad Dog' Bistany
On 1951 Mass. Charges

A Reformed Bank Bandit Searches His Soul in Print

Bank robbers
Notebooks
Final product...

http://www.aip.org/history/nbl/collections/goudsmit/index.html
Melanie J. Mueller
Archivist
mmueller@aip.org
301-209-3179