Lessons learned: Collaborating to digitise Yiddish-language collections at Cornell

Barbara Morley

Digital Archivist (retired), Kheel Center for Labor-Management Documentation & Archives at Cornell University, USA

Barbara Morley is a former digital archivist at the Kheel Center for Labor-Management Documentation & Archives at Cornell University. She has also held leadership positions in the Society of American Archivists Labor Archives Roundtable and the New York Archives Conference. She earned her MLS with a concentration in archives management from the University at Albany, State University of New York, and a Digital Archives Specialist certificate from the Society of American Archivists.

Kheel Center for Labor-Management Documentation & Archives, 227 Ives Hall, ILR School, Cornell University, Ithaca, NY 14853–3901, USA
Tel.: +1 607 255 3183; E-mail: blm2@cornell.edu

Steven Calco

Research Archivist, Kheel Center for Labor-Management Documentation & Archives at Cornell University, USA

Steven Calco is a research archivist at the Kheel Center for Labor-Management Documentation & Archives, where he provides reference, instruction and outreach to Cornell and the scholarly community at large. Steven holds a master of library and information science with a concentration in archival studies from CUNY Queens College and an MA in labour studies from the Murphy Institute, CUNY School of Professional Studies.

Kheel Center for Labor-Management Documentation & Archives, 227 Ives Hall, ILR School, Cornell University, Ithaca, NY 14853–3901, USA
Tel: +1 607 255 4318; E-mail: sc2899@cornell.edu

Elizabeth Parker

Technical Services Archivist, Kheel Center for Labor-Management Documentation & Archives at Cornell University, USA

Elizabeth Parker is the technical services archivist at the Kheel Center for Labor-Management Documentation & Archives. She serves on the Inclusive Descriptive Practices Task Force, the Collection Development Task Force on Anti-Racism and Just & Equitable Futures, and the Content Collection Policy Group for Cornell University Library's Digital Collections Portal. She holds a master of library and information science with a concentration in archival studies from CUNY Queens College, an MA in classical studies from Columbia University and a BA in history and classical studies from Colby College.

Kheel Center for Labor-Management Documentation & Archives, 227 Ives Hall, ILR School, Cornell University, Ithaca, NY 14853–3901, USA
Tel: +1 607 255 3183; E-mail: eep63@cornell.edu

Abstract   This paper describes a grant-funded collaboration between Cornell University Library, the Kheel Center for Labor-Management Documentation & Archives, Digital Consulting and Production Services, and faculty members from Cornell's Jewish Studies Program to select, digitise, describe and disseminate English and Yiddish-language records created between 1930 and 1953 by the Jewish People's Fraternal Order division of the International Workers Order. To serve researchers in the Yiddish-language community as well as English readers, Yiddish-language documents were partially translated and

transliterated to reduce the language barrier and facilitate discovery and analysis. The project required a wide variety of specialists to communicate effectively and negotiate goals, workflows and standards. As this paper discusses, responsibilities such as file and document management, metadata creation and review, and quality control are especially demanding when working with documents composed of a language and alphabet that most of the project participants do not understand. Challenges and opportunities encouraged collaborators to re-evaluate and redesign certain existing workflows, and reaffirmed the benefit of others. The online dissemination of the records has opened access to a global audience, promoting scholarly work based on the digitised records along with greater interest in related analogue collections.

KEYWORDS: Yiddish, digitisation, project management, metadata, quality control, labour archives, collaboration

INTRODUCTION

In the early 20th century, Yiddish was a vital communication medium for millions of people, including large numbers of immigrants to New York. As many of these immigrants joined labour unions, collections in the Kheel Center for Labor-Management Documentation & Archives at Cornell University include countless documents created by Yiddish speakers to educate the public and mobilise participation in cultural, economic and political activities. One such collection is the Jewish People's Fraternal Order (JPFO) — a division of the International Workers Order (IWO). Following a complex undertaking to preserve, digitise, describe and disseminate selected documents from these records, this collection is now available to a global audience. This interdisciplinary work required generous grant funding, responsive technology and insightful project design and management to succeed. Historians, archivists, conservators, metadata and digitisation specialists and systems-oriented librarians worked to bridge the disciplinary divide. Collaboration among skilled, knowledgeable stakeholders enabled them to respond to the opportunities and challenges that arose while negotiating roles, workflows, goals and standards. Project planning and implementation were complicated by the fact that only three

participants were able to read Yiddish. The resulting online collection not only supports research, teaching and learning by former JPFO members, scholars and the general public, and draws attention to the Kheel Center's related analogue collections, but also fills significant historical gaps and highlights the contributions made by underrepresented communities of Yiddish-speaking Americans that had been hidden by language barriers and historical bias against both communist-leaning and socialist-leaning organisations.

BACKGROUND

The Kheel Center is Martin P. Catherwood Library's rare and distinctive collections unit. It is part of Cornell University's ILR School, which was originally founded as the New York State School of Industrial and Labor Relations in 1949.[1] It is the largest repository documenting labour and management in North America. During the past decade, the Kheel Center has worked on several projects to increase access to Yiddish-language materials. Despite significant language barriers, these documents were selected for attention due to the wealth of materials in Yiddish at the Kheel Center and the need to preserve a language the use of which has been in steady decline since the Second World War.[2]

Curtis Lyons, the Harriet Morel Oxman Director of Catherwood Library, took special interest in digitising Yiddish-language materials during his leadership of the Kheel Center, privately describing Yiddish as the '*lingua franca* of unions' during the early 20th century. With the Center's strong collecting focus in the garment industry, many holdings include Yiddish-language materials in such collections as the International Ladies' Garment Workers' Union (ILGWU), the Amalgamated Clothing Workers of America and the International Fur and Leather Workers Union, as well as grassroots organisations including the International Workers' Order (IWO). As Central and Eastern European immigrants joined New York City's garment industry, people of different nationalities used Yiddish to communicate with each other. This contribution is demonstrated in the ILGWU's broadsides and posters used to promote labour actions, many of which were written in Yiddish.[3]

Recent scholarly and institutional endeavours to preserve Yiddish were spearheaded by the Yiddish Book Center (https://www.yiddishbookcenter.org/) and the YIVO Institute for Jewish Research (https://www.yivo.org/). Due to their efforts, the importance of the language in understanding the cultural contributions of Jewish Americans became apparent, leading to the recognition of the need for more preservation, digitisation and translation efforts by repositories with significant Yiddish holdings. Lyons saw an opportunity to work with other institutions to make the Kheel Center's Yiddish-language documents accessible to future generations.

Kathryn Dowgiewicz, then ILGWU project archivist, oversaw some Yiddish text digitisation projects, including the digitisation of the microfilmed publication *Gerechtigkeit*, the Yiddish version of the ILGWU publication *Justice* which was published from 1919 to 1969 and the

ILGWU's *The Ladies' Garment Worker*, printed in both English and Yiddish in the same publication from 1910 to 1918. As none of the project staff could read Yiddish, however, no descriptive metadata were created for that content.

The Kheel Center began to explore partnerships with institutions that had Yiddish-language materials and staff able to read the language. A collaboration with the Modern Records Centre at the University of Warwick in the UK to offer documents for crowd-sourced translation foundered,[4] as did a second discussion about presenting *Gerechtigkeit* on the National Library of Israel's digital collections portal. Unfortunately, the lack of expertise in translating Yiddish, working with volunteer labour, changing institutional priorities and technical issues, forced the Kheel Center to rethink this strategy.

As Curtis Lyons has privately commented, when it comes to successful inter-institutional collaborations, there is no substitute for two people 'clicking'. This is apparent in the Kheel Center's collaboration with the Yiddish Book Center (YBC), a cultural institution on the campus of Hampshire College dedicated to preserving the Yiddish language.

In 2015, YBC's Director of Collections Initiatives, Eitan Kensky, contacted former Kheel Center Director Cheryl Beredo to propose a joint project. The two worked well together and the YBC became a valuable ally. After Kensky left the institution, the web developer and digital projects librarian, Amber Kanner Clooney, worked with colleagues and students to create and translate metadata and to disseminate *Gerechtigkeit* on the YBC's existing portal of Yiddish-language materials.[5]

Gerechtigkeit was the first periodical ingested in the YBC portal, which previously focused exclusively on books. It was then added to the Internet Archive, providing new levels of access to the publication. As a

further service to the community, students from the YBC created metadata in Yiddish and transliterated headlines into Roman characters so users could search in both formats. Due to the software's inability to discern newspaper columns and generate useful, searchable text, however, they were unable to use optical character recognition (OCR) on the newspapers.

As of May 2021, YBC has uploaded 598 issues that have been viewed over 5,500 times, serving a community of users specifically interested in Yiddish-language materials. This project was the Kheel Center's first successful collaboration with an outside group in providing access to Yiddish-language materials. Also available from the ILR School's digital repository, DigitalCommons@ILR,[6] *Gerechtigkeit* has garnered approximately 10,000 downloads since June 2014.

When Jonathan Boyarin and Elissa Sampson, Jewish Studies professors at Cornell, became aware of the Kheel Center's rich Yiddish-language resources, they proposed a partnership to digitise material in the IWO collection highlighting the Jewish People's Fraternal Order (JPFO).

The IWO was founded as an immigrant fraternal order providing high-quality, low-cost health and burial insurance to its members. Characterising the organisation, Sampson notes that 'While the vast majority of the IWO's members — ~200,000 at its peak right after World War II — did not belong to the Communist Party of the United States of America, the IWO's politics and leadership were largely aligned with those of the Party'.[7] In 1953, due to the Cold War 'Red Scare', the organisation was dissolved by the State of New York and its records and assets seized.

In 1960, the Kheel Center acquired this collection from the Liquidation Bureau of New York. The collection includes 54 linear feet of memos, minutes, correspondence, convention proceedings, broadsides, newspapers, educational materials, political pamphlets, musical programmes and many other items in English, Yiddish and other languages.

THE ARTS AND SCIENCES GRANT

In 2012, Cornell University Library (CUL) and the College of Arts and Sciences' (A&S) Visual Resources Advisory Group[8] were allocated funds to award digitisation grants to A&S faculty and graduate students making personal, library or university resources accessible to a wider group of users. Cornell University's Grants Program for Digital Collections in Arts and Sciences (A&S grant)[9] allows for a holistic approach to digitisation and dissemination projects in consultation with Digital Consulting & Production Services (DCAPS), CUL's high-end digitisation service for the Cornell community.

Founded in 2003, DCAPS describes itself as:

> 'a virtual group that spans multiple departments within CUL, including Library Technical Services and Cornell University Library IT. Our dynamic staff includes librarians, photographers, metadata specialists, web designers, programmers, and project managers'.[10]

In 2016, Boyarin and Sampson applied for an A&S grant to digitise parts of the IWO collection, becoming the grant's principal investigators (PIs). Using their deep subject knowledge and fluency in Yiddish, they were able to provide new insights into the collection's meaning and research value. Their partnership offered deep analysis of the documents, suggesting use cases that digitisation and access could support. Sampson affirmed the scholarly value of the IWO/JPFO archives documenting the group's progressive work organising around race, ethnicity, gender, immigration, language and the arts during

the rise of anti-Fascist activism from 1930 to 1953.

PROJECT PLANNING

After the A&S grant proposal was shortlisted for funding, Barb Morley, the Kheel Center digital archivist, Dianne Dietrich, the DCAPS coordinator and central project manager, DCAPS team members and the PIs worked together to refine the project proposal, establishing clear goals and deliverables that met user needs and cultural heritage institution standards. Based on these, they defined the project's scope and objectives, created workflows, identified required skills and resources, and gauged the timeframe and budget necessary to meet project goals.

The Kheel Center's collection management database is scripted to compile folder-level container listings for each collection in Excel spreadsheets, and the PIs used the IWO/JPFO listing as the basis of their collection review, identifying candidate folders for inclusion in the project. As experienced researchers, the PIs decided which additional descriptive metadata were needed to support intellectual analysis of the material, including folder contents, document language, genre, condition, size, format and historical context, and selected high-priority folders. For those folders, the digital archivist added the folder-level filename root to the spreadsheet and student assistants wrote that alpha-numeric string on folders to aid the digitisation team (see Figure 1).[11] The spreadsheet was available to project partners in Cornell's online file sharing service, Box (https://it.cornell.edu/box).

For technical aspects of the work, the DCAPS staff reviewed selected folders, noted document types and formats, evaluated legibility, fragility and condition along with digitisation strategies, skill levels, file management needs and time requirements.

An item-level review by conservators identified documents needing treatment to facilitate digitisation and anticipated future handling. Good conservation practices determine optimal treatment for each document with respect to its digitisation and use context. If collection stewards predict an analogue master will get additional handling as an outcome of digitisation and resulting scholarly interest, conservators decide whether stabilisation is needed to facilitate heavier use in teaching, research or physical exhibits. Additional treatment can be undertaken later if uses are more demanding than anticipated.[12]

With all available information in hand, the DCAPS staff build the budget. DCAPS works on a cost-recovery model, and each project adds to an understanding of the equipment, facilities, people, knowledge and time required to do each task, reducing the difference between the estimate and actual cost of a project. Budgeting for specific project inputs and outputs is challenging. Efforts to design and build sustainable collections require planning up front, but also continued monitoring and maintenance over many years. These budget elements must try to account for the dynamic, technical array of design and functional options, as well as the changing preservation and access environment, making real-world expenditure forecasting very difficult (see Figure 2).

Based on this thorough analysis by an experienced team, the Visual Resources Advisory Committee decided to fund the IWO/JPFO project due to the collection's high research value and unique scope. This formalised the first collaboration between the Kheel Center and faculty from the Jewish Studies Program and between those faculty members and DCAPS. It built on the digital archivist's 15 years collaborating with DCAPS team members on digitisation work.

5276 Box	5276 Folder	Folder name	Copyright notices	Description	Contents notes, brittle condition	File name root
29	1	Kahn, Albert E. - President	no copyright	Correspondence re situation in Poland; Jews in Palestine; form letters; memo, "Statement on Second Session of World Jewish Congress, 1948, to Executive Committee Members."	**Much of box 29 is BRITTLE.** Includes געשיכטע פון דער פראטערנאלער באוועגונג [AUTHOR INSCRIBED]. Nyu York; Internatsionaler Arbeter Ordn, 1936. Original Cloth. 8vo. 287 pages. 23 cm. First edition. 'On the History of the Fraternal Movement'; an early organizational history of the International Workers Order, with biographies of prominent personalities, including photographs, as well as charts detailing the growth and work of the organization. Written and inscribed by the first General Secretary of the IWO Rubin Saltzman.	5276b29f01_
36	2	Cultural Department, Poems and Songs	no copyright	A Song to the Order (A Gezung Tzum Ordn fun Shmole Davidyan?) Brider un Shvester fun Shtut un Fabrik; Juniorama about Poems for Victory by Eric Lucas describing the kids from 3 camps (including Kinderland writing "Victory Verses for Young American;" Bi-Lingual Program of Twenty-Fifth Jubilee mid-Winter Concert-Pageant, Jewish Folk Chorus, Prof. S.G. Braslavsky, Conductor, Jan 23, 1949, New England Mutual Hall, Boston with 3 Pictures by Rubinstein, and a piece by Moussorgsky and then a Pageant of 50 years of Jewish Life (listing each song), A Singing People Lives Forever, Fifty Years of Jewish Life Mirrored in Folklore and Song by M Repook...	Mainly Yiddish. Doesn't miss a beat on the affective buttons and fully integrates the USSR propaganda. Encore: HaTikvah Recapitulates JPFO politics quite nicely at this time period in which they are trying to reposition themselves as part of a shared war effort and the future of the Jewish people.	5276b36f02_
53	18	R. Saltzman correspondence with Poland	no copyright	In Yiddish. Re: relief efforts in Poland, e.g. telegram asking for 100 grams streptomicin. Handwritten letter Dec 4, 1946. Radiogram from Warsaw Sept 15, 1947 in English wishing the JPFO a Happy New Year (Rosh HaShana). Letter from Central Committee of Jews in Poland (July 10, 1947) followed by hand-written and type correspondence from that July. Conditions are bad. Two other radiograms asking for antibiotics immediately. Invitation to a JPFO exhibit (May 8 1949) "The Jew in Poland: from Ruins to a New Life" at ACA Gallery, 63 E 57th Street NY (shows Albert E. Kahn as President and Rubin Saltzman as General Secretary)	Handwritten as well as typed materials. Very brittle. Radiograms have good graphics. The material from Poland shows post-war conditions and JPFO fund raising and relief efforts. Exhibit invite is also interesting politically as well as culturally	5276b53f18_
53	7		no copyright	IKUF: Alveltlekher Yidisher Kultur-Farband "IKUF" (World Jewish Culture Association), national committee of the IKUF in America: Kalmen (Kalman) Marmor (Mermer), Z. Weinper, Paul (Pesach) Novick, Nachman Meisel, R. Zaltsman, Itche Goldberg, [Rabbi? Bik], Mahler, Kleinberg, B. (B.Z.) Goldberg (Benjamin Waife, Sholem Aleichem's son-in-law), Moishe Katz, Ahron Gudelman (Aaron Gidelman)	IKUF logo in green on letterhead is quite nice. Very brittle letters in Yiddish from Paris, New York, and Poland on the founding of the organization. Polish letter is handwritten. Box 53 Folder 7 has some founding documents from Paris, NY, and Poland. A few letters	5276b53f07_

Figure 1: Excerpt of exported folder list from collection 5276, with document descriptions and comments from the PIs

PROJECT MANAGEMENT	Hours	Cost per hour	Subtotal	Notes
First year	100	$65	$6,500	Minimum flat fee: 100 hours at $65/hr
Additional years		$65	—	
			$6,500	

PREPARATION	Hours	Cost per hour	Subtotal	Notes
Conservation Assessment	3	$65	$195	
Acquisition of Physical Material	10	$65	$650	
Conservation Treatment	30	$65	$1,950	
			$2,795	

RIGHTS MANAGEMENT	Hours	Cost per hour	Subtotal	Notes
Copyright and IPR Consulting	20	$85	**$1,700**	Ask Amy Dygert, Director of Copyright

METADATA	Hours	Cost per hour	Subtotal	Notes
Design	20	$80	$1,600	shared shelf
Implementation	125	$55	$6,875	
			$8,475	

STORAGE AND BACKUP	Size	Cost per GB	Subtotal	Notes
Maintenance				
Preservation				

DEVELOPMENT	Hours	Cost per hour	Subtotal	Notes
Portal/Design and Development	40	$85	$3,400	Delivery is slated for the Hydra-based digital collection portal at CUL
Portal/Backend Programming	20	$85	$1,700	
			$5,100	

Photography	Quantity	Hours	Cost per hour	Subtotal	Notes
Capture	10,000	395	$65	$25,675	~4100 flatbed, rgb 300–400 dpi, 22/hour equals 186 hours =$12,090 at current $65/hour
Image Processing		55	$65	$3,575	~1000 flatbed, rgb 600 dpi, 12/hour equals, 83.3 hours = $ 5,414 at current $65/hour
Basic Structuring		40	$65	$2,600	
Structuring/Tagging		20	$65	$1,300	~ 79 large fragile zeutschel or Nikon at ~$7 ea = $ 553 at current $65/hour
Quality Control		**20**	**$65**	**$1,300**	~880 difficult to handle zeutschel at ~$3.50 ea = $3,080 at current $65/hour
				$34,450	~2514 normal Zeutschel scans at ~ $1.50 ea = $3771 ~211 sheetfed (bur may need fb) ~.20ea sf, or ~$3 sf = $ 42 - $633

Figure 2: DCAPS budget for the IWO Digitisation Project

Audio Visual	Quantity	Hours	Cost per hour	Subtotal	Notes
Capture					
Processing					
Metadata/Derivative Creation					
Quality Control					

PROCESSING	Quantity	Hours	Cost per hour	Subtotal	Notes
OCR		15	$80	$1,200	
Data Remediation		—	—	—	
Shared Shelf		2	$80	$160	
				$1,360	

MISCELLANEOUS	Quantity	Hours	Cost per hour	Subtotal	Notes

| TOTAL ESTIMATE | | | | **$60,380** | |

Figure 2:   (*continued*)

TECHNICAL CONSIDERATIONS

Two of the biggest considerations and hurdles to any digitisation project are the technological underpinnings and the comfort of stakeholders with these tools. CUL is fortunate to have specialists to oversee digitisation, file management and management of the Library's Digital Collections Portal.[13] The portal is the preferred delivery system used by CUL to host and provide access to digital assets. In the IWO/JPFO project, one goal was to have the selected materials uploaded to an online dissemination platform with descriptive metadata and a landing page to contextualise the curated materials. To make the collection accessible, the Yiddish and English components of the digitised materials had to be taken into consideration. Although CUL uses HathiTrust as a valuable portal for other documents, it was not chosen because of the strict criteria for mixed-language collections.

Artstor's JSTOR Forum was selected as the metadata creation tool because it is an open source program developed in concert with its user community. While the curated selections of digitised materials have both item-level metadata and a collection landing page, the portal was not designed to host an online exhibit as the PIs initially hoped. The distinction between a portal for discovery and an exhibit space frequently causes friction in digitisation and dissemination projects.

Providing full-text access to multilingual collections in a digital environment presented additional challenges. To support full-text searching in documents, file management specialist Mira Basara used Abbyy Finereader 15 OCR processing software (https://pdf.abbyy.com/), giving Yiddish-language readers a way to do keyword searching with Hebrew characters.

OCR technology has its own hurdles. These include interpreting columns in text, recognising typefaces that may not be clear to the algorithm, reading originals of mixed quality which may not render readable text, and the inability to interpret most handwritten documents. Despite these issues, the PIs felt it was important to provide a basic level of accessibility for the Yiddish-reading community. Time limitations for non-Roman character training and processing led CUL's file management specialist to focus additional energy on collaborating with digitisation specialists to ensure high-quality captures for potentially problematic documents.

PROJECT IMPLEMENTATION AND MANAGEMENT

Trusting relationships are essential for effective collaboration[14–16] and are enhanced by shared experiences with effective problem solving. Although difficult compromises can leave residual stress points, for a project to succeed there must be good communication, and the plan must have well-defined, achievable project scope and goals that match available resources, with clear participant roles.

Catherwood Library Director, Curtis Lyons, emphasised the value of face-to-face discussions to forge strong, personal relationships among partners. Grant kick-off meetings allow large, interdepartmental teams to contextualise the project from multiple perspectives while clarifying one another's roles and available resources. Frank discussions help identify unforeseen opportunities and potential stumbling blocks and support effective problem-solving based on mutual understanding and trust. Discussing responsibilities, expectations, capacities and boundaries helps avoid setting unachievable goals and minimises project scope creep.

Proposals that are unrealistically large or technically demanding need to be scaled

back so they can be accomplished within the resource constraints of the grant.[17] Years of experience have also made it clear that some elements in a project may need fine-tuning or full revision during implementation.

As acknowledged by the former Kheel Center Director, projects must have sufficient time to go 'off the rails'.[18] When a problem arises, participants need time to recognise it, diagnose it, develop and evaluate options, and implement them, along with any cascading adjustments that might be indicated. DCAPS has weekly project management meetings to assess progress, refine workflows and solve problems, and uses Confluence, a project management and documentation tool, to connect teams and departments from designing workflows and estimating budgets, through documenting decisions and tracking progress.[19] DCAPS also uses Zoho[20] to track inputs, outputs and billing.

Managing resources is especially challenging with new types of grant projects as, after implementation begins, unanticipated variables can arise. Each component, human or technical, must be available at the right time for projects to move through phases smoothly, employing sequenced assets efficiently. This remains the case at the tail-end of a project when file delivery, quality control (QC) and rescans or metadata clean-up occur, after many participants have reallocated their time and resources to other work.

Unspecified assumptions, resource limitations and conflicting priorities may delay anticipated progress or block success. Three sources of difficulty were noteworthy, costing unbudgeted time and money for the Kheel Center, PIs and DCAPS. These included planning for labelling analogue masters with their corresponding file names, creating metadata for Yiddish-language documents and QC for surrogate files and metadata.

For typical Kheel Center digitisation projects, student assistants provide essential labour, contributing both subject knowledge

and skills. Student assistants usually work with the digital archivist to review the collection and create item-level metadata. They pencil the assigned file name on the corresponding document before digitisation to ensure that the analogue master, associated metadata and digital surrogate are assigned matching identifiers that conform to the Kheel Center's file-naming convention.

The A&S grant programme includes digitisation, file management and item-level metadata creation for documents previously described at the folder level only, so the digital archivist did not follow the typical workflows. The digital archivist incorrectly assumed that the digitisation team would label analogue masters as they assigned file names during digitisation. When analogue masters, files and metadata were delivered to the Kheel Center, the planning oversight became obvious. Digital collections student assistants spent many days comparing analogue masters with digitised content and labelling originals with the file name as part of the QC process — a costly oversight even at student wage rates.

Metadata creation was also challenging. Because A&S grant PIs have deep subject expertise, they or their research assistants typically create metadata for funded projects. During the application and planning phase of the grant, the PIs were not aware that they would be expected to create the metadata and the DCAPS staff did not realise the PIs could not allocate time to do that work. When this misunderstanding was discovered, DCAPS hired a part-time, temporary, metadata assistant who reads Yiddish.

This large digitisation project was one of the first projects handed off to Jasmine Burns, the newly hired visual resources metadata librarian. Despite her strong background in metadata creation and management, this would be the first time that Burns had been responsible for supervising a metadata assistant working in another language — all while still learning about CUL and her other responsibilities.

Because the creation of highly detailed metadata took longer than the budgeted six months, a second metadata assistant had to be hired to create metadata for English-language documents. The assistants lacked relevant historical knowledge and required extensive training and supervision, creating unanticipated demand on the metadata librarian's time. A further — appropriate, but unbudgeted — expense was the university's requirement that temporary employees working longer than six months should receive standard employee benefits.

Historians describing similar materials for the scholarly community emphasise historical significance, underscoring the importance of detailed descriptive metadata. The metadata records created by the metadata assistants did not meet the expectations of the PIs, who then spent many days correcting and enhancing the initial entries with more extensive summaries in the description field (Figures 3 and 4). This level of detail was more typical of online exhibition curation (or what one project participant called 'research data'), deviating from common metadata field use and exceeding customary, item-level allowances for library metadata.[21] The supervisors' lack of Yiddish-language expertise made it difficult to evaluate metadata as they were being created and to identify issues in their infancy. Were this massive investment of time in collection description undertaken as a regular practice, it would result in richer item-level records but would also hinder the institution's ability to describe, present and maintain large collections.

Detailed work was done to translate and transliterate each title, which the PIs felt was important so English-language researchers could acquire a cursory understanding of the documents. This required two alphabets — Romanised Yiddish and Hebrew script — in the title field to help meet accessibility goals and transliteration standards. PIs used transliteration guidelines developed by the YIVO Institute for Jewish Research — the

Figure 3: JSTOR Forum description field, work types, Note field

Subject:
World War II- War Effort, Fundraising
Soviet Union
Culture Front- Publishing
Intelligentsia
Jewish Left
Communism
Jewish Anti-Fascist Committee of the USSR (JAFC)
Wartime Jewish Unity
Fellow Travellers
Nazism and Fascism
Fellow Travellers
Jewish Lodges

Description:
Fundraising leaflet with Einstein's address from the October 25, 1942 dinner in his honor at the Hotel Commodore in New York. The Jewish Council for Russian War Relief, Inc. is part of Russian War Relief, Inc. Einstein praises Russia's scientific, technological and economic progress, reviews Russia's pre-war foreign policy, and offers some words of mitigation for the lack of internal political freedom. Implies the U.S.S.R is doing a better job of treating minority nationalities equally, which he calls, "a fact of peculiar and decisive importance for us Jews". Finally he turns to the crux of the argument, that if Russia fell, England and America would be "in a very bad way." Drawing by Morris J. Kallem.

Notes:
The Jewish Council for Russian War Relief, Inc. is part of Russian War Relief, Inc..
The Jewish People's Fraternal Order was the largest 'national' section of the International Workers Order (IWO) which focused on cultural awareness and celebration, mutual support especially in health insurance coverage, and anti-fascist activities. The IWO also gave particular emphasis to supporting the rights and interests of African Americans. Documents include language and representations which comprise the historical record and should not be interpreted to mean that Cornell University or its staff endorse or approve of negative representations or stereotypes presented.

Figure 4: Portal front end, displaying the metadata as seen by users

framework predominantly implemented by publishers and scholars.[22] To ensure consistency with YIVO standards, the transliterations made by assistants were saved in a shared 'review' folder, which the PIs edited as needed. The reviewed files were then moved to an 'approved' folder from which the metadata librarian was able to batch upload the revised entries. The PIs often used in–depth analysis of related documents to make decisions on complex translation and transcription issues. For example, Professor Sampson discussed the 'ambiguities in the use of a common Yiddish word like *Chaver (Khaver)*, which can mean brother, comrade, friend, member, etc, and can be spelled with Hebrew orthography or in a few instances with Russian Yiddish

orthography'.[23] The PIs also made sure to include original transliteration and English translations from the source documents themselves in the metadata, even where standardised Yiddish had also been entered into the fields. This was done to respect the grammar and orthography of the creators of the documents.

Yiddish names presented unique hurdles. In the metadata profile, the names are given in Hebrew script, Romanised Yiddish and English (Figure 5). When these documents were created, there was no standardised system to transliterate Yiddish and there was intense societal pressure to anglicise names. This led to inconsistency in the names used by prominent figures and meant that every name authority used in the metadata

Figure 5: JSTOR Forum backend, showing metadata for correspondents; the 'Working' tab shows name authority fields in English and Yiddish for review

had three distinct entries and needed to be checked by the Yiddish metadata assistant, the metadata librarian and the PIs. Mindful of unintended changes to the approved entries, the metadata librarian customised the JSTOR Forum interface, blocking fields so the Yiddish and transliterated name authorities could be checked against English name authorities without risking alterations to already approved metadata.

Labour issues surrounding metadata creation and enhancement have a large impact on the project team and the final product. It is very easy for a challenging project to consume more time and resources than allocated, potentially short-changing other work.[24] As the visual resources metadata librarian pointed out, description is an iterative process that requires a mechanism to allow collection stewards to create necessary changes with minimal impact on metadata librarians.

Toward the end of the project, Professor Sampson did two rounds of enhancements in the JSTOR Forum that were checked and uploaded by the metadata librarian. The PIs later had opportunities to edit metadata in spreadsheets, making it easier for the metadata librarian to check any amendments. As noted by Sampson and Boyarin, requiring PIs to create the metadata limits the number of faculty and graduate students who can propose projects. Most lack time to create metadata while meeting other departmental requirements for promotion or graduation.[25]

QUALITY CONTROL

QC is labour-intensive, and much more difficult than one might think, requiring sustained, resolute attention to detail. Full QC is a time-consuming step that few projects can afford but most could benefit from. Many plans allocate QC resources to one-fifth or one-tenth of the content, supplementing notes made by digitisation staff documenting faded, torn, skewed,

duplicate or missing pages that cannot be captured more clearly or completely.

For Kheel Center digitisation projects done in collaboration with DCAPS, DCAPS staff members do the first round of QC and Kheel Center digital collections student assistants do a second, comprehensive review. This ensures that every analogue master has a corresponding digital surrogate; that the file name for each analogue master is correct, pencilled on the back of the master, and that it matches the surrogate files; that each file is complete and correctly assembled when compared with the analogue master; that there are preservation TIFFs and access PDFs that have undergone OCR scanning; that the metadata describe the document; and that the correct document image is paired with the metadata record. Any corrections that the student assistant or digital archivist is unable to make are communicated to DCAPS using the shared spreadsheet. DCAPS staff note their questions, suggestions or corrections in the spreadsheet, ensuring that each participant knows the status of each issue and the project as a whole.

Even with a detailed workflow to help reviewers stay on track, QC for large sets of text is arduous, especially if reviewers do not read the language's script. Despite not reading Yiddish, one particular digital collections student assistant, Serina Moheed, was frequently assigned QC work because of her sustained attention and ability to identify errors. Moheed likened her process of reviewing Yiddish-language documents to the examination of photographs: along with her colleagues, she would look for right-justified content and occasional Arabic numerals to establish orientation; consult Yiddish-learner websites and lists of days and months to establish order and proper assembly; and compare graphics and the positions of advertisements to check for duplication. Moheed and the DCAPS digitisation specialists recommended that, rather than making QC the last stage of the

project, it should begin much earlier and be done in small batches to reduce the intensity of the work and improve success rates.

Interestingly, the clean-up portion of the project posed some of the most significant questions. What does reasonable and responsible QC look like for digitisation and metadata creation? Does ability to read the language affect the percentage of QC work undertaken? Who has the knowledge needed to recognise the meaning of errors and systematically explore other records that might have related issues? Who has the requisite attention span to make detailed comparisons, and what conditions would support their success? How can timely and effective QC and problem remediation be ensured when new projects have been undertaken and staff time has been budgeted elsewhere?

OUTCOMES

At the end of a project, a wrap-up meeting allows collaborators to discuss what went well, what could be improved, and whether each unit achieved its goals. It facilitates holistic analysis of each participant's contribution to the grant and provides a chance to celebrate the collaboration. The IWO/JPFO project had a long tail of metadata creation and enhancement, QC and file and metadata correction, including an unusual delay due to the COVID-19 pandemic. It is only drawing to a close now in 2021.

Challenges and opportunities during the IWO/JPFO digitisation project allowed collaborators to re-evaluate and redesign some long-standing divisions of labour, task assignments, tools and workflows, and reinforced the benefit of others. The head of Conservation realigned assignments to take better advantage of available staff with needed expertise. The Digitisation and Conservation Services Director highlighted the need for more project

managers and has since assigned one person to manage A&S grants, and added project management to two digitisation specialists' responsibilities. The DCAPS coordinator noted that this project influenced the way she incorporates long-term sustainability and ongoing stewardship with library partners during project planning.[26]

Strategically important outcomes included insights into the most successful ways to collaborate with a wide variety of partners and new understandings regarding the best ways to share information and communicate the goals and expectations of different communities of practice. *Ad hoc* workflows built on interpersonal relationships are often weak and require adjustment as they encounter conflicting tools and workflows. Jasmine Burns reiterated the importance of well-defined goals, setting clear expectations, defining deliverables and timetables, and formalising these in clear documentation.[27] These decisions provide a solid foundation for detailed project planning, progress assessment and tactical adjustments, if needed. This shared understanding can be expressed in a memorandum of understanding or collaboration statement, a written agreement describing the goals and products of the collaboration, naming the partners and contributions each will make, and indicating the timeline on which each aspect of the work will be accomplished.

Since working on this project, the metadata librarian has created a robust and structured template for metadata creation in JSTOR Forum. She has also used JSTOR tools to create linked data menus using subject headings from authorised vocabularies. This increases consistency across collections and allows for greater discoverability as the metadata are pushed to various hosting platforms. This also creates an easier interface for those who find it difficult to use the full interface effectively and has allowed for greater consistency among digital collections' metadata.

For the Kheel Center, this project provides a roadmap for future work describing and providing access to multilingual and non-Roman script collections. With the globalisation of the supply chain, documenting labour issues inherently includes the ability to describe and make available records in a wealth of languages and orthographies. There are numerous gaps in the archival record surrounding the lives and contributions of marginalised workers. The successful description and dissemination of the Yiddish-language records offers guidance for future collaborations to make such materials available.

Thanks to the creative design and programming efforts of the DCAPS web team working with the PIs, the digitised materials went live on CUL's Digital Collections Portal in May 2018 (Figure 6). The IWO/JPFO digital collection has widened access to information about the Yiddish Left and given scholars a new avenue to study Yiddish language and heritage. In addition to the rich document description, the landing page features a 3,700-word essay, 'Fellow Travelers: From Popular Front to Cold War', written by Sampson and Boyarin, exploring the IWO/JPFO digital collection, detailing the history of the organisation and highlighting areas of potential interest for scholars.[28] It also includes a timeline and bibliography to help contextualise the materials.

Sampson, Boyarin, a group of Jewish Studies scholars and Kheel Center archivists worked together to organise a week-long virtual conference, '*Di Linke*: The Yiddish Immigrant Left from Popular Front to Cold War' in December 2020, examining the complex history of the IWO/JPFO. Conference presentations were based on digitised materials from the IWO/JPFO digital collection.[29] Six hundred people from around the world were in attendance, including former JPFO members, interdisciplinary scholars, digital humanists, researchers, students, activists and the general public. Re-energised interest in these materials encouraged project participants to apply for a second round of funding to add to the digitised collection.

Other important outcomes include numerous related donations to the Kheel Center, an oral history programme established with former IWO/JPFO members, and increased interest from scholars and the general public in accessing both digital and analogue IWO materials. Research archivist Steven Calco is now able to provide patrons with genealogical assistance using Yiddish-language resources. In 2018, he connected a Holocaust survivor and war orphan with articles and photographs from *Gerechtigkeit*, as well as a digitised video of her participation in the 1950 ILGWU convention as a cultural ambassador from war-torn Europe.[30]

CONCLUSION

In the past, the availability of Yiddish-language records in high-use Kheel Center collections drew the attention of archivists and scholars, but very few could read the language. Contributions by content specialists, Yiddish-readers and technical specialists has resulted in not only preserved, digitised documents but also rich metadata transliterating and translating essential elements of the documents, enabling discovery and access by both Yiddish and English-language readers.

The online resources created during this collaborative project foster community engagement, project development and scholarly inquiry in the Kheel Center's Yiddish-language collections. As Rokhl Kafrissen notes:

> 'It's hard to describe just how fascinating this material is, and how important is the work being done in its archive. The IWO and JPFO have been so thoroughly scrubbed from Jewish American history, and American history, that most of us simply have no reference point.'[31]

CORNELL UNIVERSITY
Library

Digital Collections

International Workers Order (IWO) and Jewish People's Fraternal Order (JPFO)

Fellow Travelers: From Popular Front to Cold War. Selections from the ILR School Catherwood Library Archives of the Yiddish Immigrant Left

Jewish People's Fraternal Order (U.S.)

ייִדישער פֿראָטערנאַלער פֿאָלקס-אָרדן

JPFO, א.פ.פ.א.

International Workers Order

אינטערנאַציאָנאַלער אַרבעטער אָרדן

IWO, א.א.א.

Browse by set | browse full collection

Search... All Fields ⌄ 🔍

View timeline of IWO and JPFO (PDF)

View bibliography (PDF)

This bilingual project conserves, digitizes, and curates* a portion of Cornell's International Workers Order (IWO) archives and most especially those of its Jewish division, known as the Jewish People's Fraternal Order (JPFO). The IWO was founded in 1930 as an immigrant fraternal order that provided high-quality, low-cost health and burial insurance and other benefits for members. The origins of the IWO / JPFO arose from a decade of splits (1920–1930) concerning the U.S.S.R., the Bolshevik Revolution and Communism that consumed the Jewish Federation of Socialists and the Arbeter Ring (Workmen's Circle) groups associated with Eugene V. Debs' Socialist Party. While the vast majority of the IWO's members—~200,000 at its peak right after World War II—did not belong to the Communist Party of the United States of America, the IWO's politics and leadership were largely aligned with those of the Party. For those familiar with the Yiddish speaking immigrant Left, this group was often referred to as "Di Linke" or the Left. The IWO was legally disbanded in 1953 due to the Cold War "Red Scare." This closure followed on a famous and unprecedented court case, in connection with which the organization's insurance funds and records were seized by New York State's Insurance Department. The presence of a substantial portion of the IWO Records (#5276) in the Kheel Center at Cornell's ILR School is a direct result of that seizure.

Figure 6: Excerpt from collection landing page

As with so many large projects, undertaking the documents' conservation, digitisation, description and dissemination required grant funding that was available, along with project expertise, as part of Cornell University's Grants Program for Digital Collections in Arts and Sciences. PIs from Cornell's Jewish Studies Program curated document selections and prepared the in-depth landing page essay. Relying on extensive communication and shared project documentation, Cornell University Library's DCAPS specialists, metadata creators and the Kheel Center's digital archivist identified challenges, designed and managed workflows, and discerned and resolved problems. This highly collaborative effort culminated in a successful project to make Yiddish and English-language documents from the JPFO available on CUL's Digital Collections Portal to a worldwide audience including former JPFO members, students, scholars and the general public. The project also had the unexpected result of adding an entirely new layer of documentation to the Kheel Center's collections, keeping the story of the IWO/JPFO vibrant and filling gaps in the archival and historical record.

References

1. Kheel Center (2021) 'Kheel Center for Labor-Management Documentation & Archives', available at: https://catherwood.library.cornell.edu/kheel/about/ (accessed 6th May, 2021).
2. Roses, J. (1996) 'A dead language, "Yiddish lives"', *New York Times*, 7th July, available at: https://www.nytimes.com/1996/07/07/magazine/a-dead-language-yiddish-lives.html (accessed 6th May, 2021)
3. Katz, D. (2013) 'All Together Different: Yiddish Socialists, Garment Workers, and the Labor Roots of Multiculturalism', NYU Press, New York, NY.
4. University of Warwick (2014) 'Appeal for Yiddish speakers to help translate UK and USA archives', available at: https://warwick.ac.uk/newsandevents/pressreleases/yiddish_speakers_to/ (accessed 6th May, 2021).
5. The Yiddish Book Center (2021) 'Gerechtigkeit (Justice)', available at: https://www.yiddishbookcenter.org/collections/gerechtigkeit-justice (accessed 6th May, 2021).
6. DigitalCollections@ILR (2021) 'Gerechtigkeit (Yiddish Justice)', available at: https://ecommons.cornell.edu/handle/1813/72812 (accessed 6th May, 2021).
7. Cornell University Digital Collections (2021) 'International Workers Order (IWO) and Jewish People's Fraternal Order (JPFO)', available at: https://digital.library.cornell.edu/collections/iwo-jpfo (accessed 6th May, 2021).
8. Digital Consulting & Production Services (2021) 'Visual Resources Advisory Group', available at: https://dcaps.library.cornell.edu/grants/as/advisory (accessed 6th May, 2021).
9. Digital Consulting & Production Services (2021) 'Grants Program for Digital Collections in Arts and Sciences', available at: https://dcaps.library.cornell.edu/grants/as (accessed 6th May, 2021).
10. Digital Consulting & Production Services (2021) 'Welcome to DCAPS', available at: https://dcaps.library.cornell.edu/ (accessed 6th May, 2021).
11. Kheel Center (2021) '#5276 International Workers Order (IWO) Records Finding Aid Guide', available at: https://rmc.library.cornell.edu/EAD/htmldocs/KCL05276.html (accessed 3rd August, 2021).
12. Michele Hamill, personal communication, February 2021.
13. Cornell University Digital Collections (2021) 'About Cornell Digital Collections', available at: https://digital.library.cornell.edu/about (accessed 6th May, 2021).
14. Dietrich, P., Eskerod, P., Dalcher, D. and Sandhawalia, B. (2010) 'The dynamics of collaboration in multi-partner projects', *Project Management Journal*, Vol. 41, No. 4, pp. 59–78.
15. San Martín-Rodríguez, L., Beaulieu, M.-D., D'Amour, D. and Ferrada-Videla, M. (2005) 'The determinants of successful collaboration: A review of theoretical and empirical studies', *Journal of Interprofessional Care*, Vol. 19, Suppl. 1, pp. 132–147.
16. Bond-Barnard, T.J., Fletcher, L. and Steyn, H. (2018) 'Linking trust and collaboration in project teams to

project management success', *International Journal of Managing Projects in Business*, Vol. 11, No. 2, pp. 432–457.

17. Dianne Dietrich, personal communication, December 2020.
18. Cheryl Beredo, personal communication, January 2021.
19. Atlassian. (2021) 'Accomplish more together', available at: https://www.atlassian.com/software/confluence
20. Zoho. (2021) 'Your life's work, powered by our life's work', available at: https://www.zoho.com/
21. Baca, M., Harpring, P., Lanzi, E., McRae, L. and Whiteside, A. (2020) 'Description', in 'Cataloging Cultural Objects: A Guide to Describing Cultural Works and their Images', Chapter 8, available at: http://vraweb.org/wp-content/uploads/2020/04 /CatalogingCulturalObjectsFullv2.pdf (accessed 2nd August, 2021).
22. Yiddish Wit (2021) 'YIVO Transliteration Chart and Yiddish Spelling Using YIVO Transliterations', available at: https://www.yiddishwit.com/transliteration.html (accessed 6th June, 2021).
23. Elissa Sampson, personal communication, June 2021.
24. Jasmine Burns, personal communication, December 2020.
25. Elissa Sampson and Jonathan Boyarin, personal communication, December 2020.
26. *Ibid.,* ref. 17 above.
27. *Ibid.,* ref. 24 above.
28. Cornell University Digital Collections, ref. 7 above.
29. Cornell Cast (2021) 'Di Linke: the Yiddish Immigrant Left from Popular Front to Cold War Conference Videos', available at: https://www.cornell.edu/video/playlist/di-linke-conference (accessed 6th June, 2021).
30. Beduya, J. (2020) 'Family ties abound in Cornell University Library archives', *Ezra Magazine*, 26th October, available at: https://news.cornell.edu/stories/2020/10/family-ties-abound-cornell-university-library-archives (accessed 6th May, 2021).
31. Kafrissen, R. (2021) 'Rokhl's Golden City: what a dog named Labzik can teach us about communism', *Tablet*, 30th April, available at: https://www.tabletmag.com/sections/community/articles/puppy-politics-communism (accessed 6th May, 2021).