The Robotics Project

CARNEGIE MELLON UNIVERSITY LIBRARIES

Carnegie Mellon University

March 2023

MULTIMODAL ARCHIVES

A Toolkit for Collecting Robotics and Other Complex Material in a Research Ecosystem

Brian Mathews, Katherine Barbera, Kathleen Donahoe, and Lauren Herckis

http://www.library.cmu.edu/robotarchive

Image: Jodi Forlizzi Collection, Carnegie Mellon University Archives

ACKNOWLEDGEMENTS

TEAM MEMBERS

- **Brian Mathews, PI**
- **Katherine Barbera, Co-PI**
- **Kathleen Donahoe, Processing Archivist**
- **Lauren Herckis**
- **Julia Corrin**
- **Andrew Meade McGee**

CONTRIBUTORS

The team would like to thank our colleagues who contributed to Phase 1 of The Robotics Project by providing technical expertise and project support: **Heidi Bartlett, David Bernabo, Angelina Crowe, Chris Kellen, Jonathan Kiritharan, Matthew Lincoln, Ann Marie Mesco, Joseph Mesco, Ken Rose, and Morgan Walbert.** This toolkit would not be possible without their efforts. We are also grateful to participants in the project's ethnographic study.

CONSULTANTS

Our team worked with consultants in a variety of disciplines to create this toolkit. **Brittany Cox, Mary Wilcop, and Rhonda Wozniak** identified conservation plans and approaches for robotics material, **Small Data Industries** provided recommendations for digital preservation and infrastructure, and **Meredith Evans** helped us craft inclusive, community-centered collection development strategies.

FUNDING & SUPPORT

THANK YOU TO OUR DONORS AND COLLEAGUES

We gratefully acknowledge everyone who supported Phase 1 of The Robotics Project. Their contributions brought this toolkit into being. A complete account of this community is not possible, but we offer this list in alphabetical order.

Mr. Bradley Allen and Mrs. Rebecca Allen
Argo AI
Christopher G. Atkeson
Ian Beitler
Jenny Belardi
Katie Coan
Catherine Copetas
Emily Davis
Blair Dunckel
Jodi Forlizzi
Gretchen Graff
Jason Griffey
Alan Guisewite
Hannah Gunderman
Dr. Satish Gupta and Ms. Sharon Edwards
Jean Harpley
Martial Hebert
Jessica Hodgins
The Roy A. Hunt Foundation
Dr. David R. Jefferson
Crystal Johnson
Matthew Johnson-Roberson
Dr. Clinton W. Kelly, III
Halil Kulluk
Samuel Lemley
Joan Lippincott
Shannon Mominee
Jon McIntire
Rikk Mulligan
Near Earth Autonomy, Inc
Leah Picker
Joelle Pitts
Mark Power
Shannon Riffe
Robotics Institute
School of Computer Science
Daniel Siewiorek
Emma Slayton
Brian Staszel
Tony Stentz
Dr. Benjamin Stephens
Debra Tobin
Keith Webster
Sonia Wellington
Red Whittaker
Manuela Veloso

This toolkit and Phase 1 of The Robotics Project were made possible by the Alfred P. Sloan Foundation and other generous donors.

ALFRED P. SLOAN FOUNDATION

ABOUT THE AUTHORS

Brian Mathews

Brian Mathews is the Associate Dean for Research & Innovation at the Carnegie Mellon University Libraries. He is interested in the intersection of creativity, interdisciplinarity, and knowledge management. He has previously held library leadership roles at Virginia Tech, Georgia Tech, and UC Santa Barbara.

Katherine Barbera

Katherine Barbera is an archivist and oral historian interested in multimodal collection and public programming strategies for archives and museums. At Carnegie Mellon University Libraries, she is the founder and director of the Oral History Program, an archivist in the University Archives, and lead archivist for The Robotics Project.

Kathleen Donahoe

Kathleen Donahoe is the Robot Archive Processing Archivist at Carnegie Mellon University Libraries. She has spent nearly ten years processing collections in university archives settings and focuses on outreach and engagement practices for libraries and archives. A native Pittsburgher, she is interested in the history of the city's industrial transformation and its current status as a robotics capital of the world.

Lauren Herckis

Lauren Herckis Ph.D. is Simon Initiative Library Faculty and holds appointments in the Carnegie Mellon University Libraries and the Human-Computer Interaction Institute. Her research at the intersection of campus culture, technological change, and effective teaching is supported by the California Education Learning Labs, the Global Learning Council, and the National Science Foundation. Current projects explore barriers to the use of evidence-based tools and practices, identify community-engaged, data-driven implementation strategies for technology-enhanced learning tools, and help leaders in higher education develop resilient postsecondary infrastructure.

TABLE OF CONTENTS

1. Introduction 6

2. Approach 10

3. Findings 28

4. The ROBOTS Conversation Guide 37

5. Closing Considerations 49

Appendix 53

Bibliography 55

Questions?

Visit our website
http://library.cmu.edu/robot-archive

1. INTRODUCTION

OVERVIEW

Robotics presents a significant challenge for archivists and information professionals. The field produces a large volume of multimodal products and by-products such as documentation, code, machines, and prototypes that depend heavily on in situ application and shift rapidly with changing technologies and research trends. This is compounded by the sheer complexity of the social networks, group processes, and interpersonal dynamics that drive research and development and take place in the context of multiple and interconnected teams, laboratories, communities, and institutions across the globe.

To explore this interconnected landscape, in 2019 Carnegie Mellon University launched **The Robotics Project**, a multi-phase collaboration between the University Libraries and the School of Computer Science. Our interdisciplinary team is investigating the research ecosystem of robotics—drawing on wide-ranging expertise emerging in modern libraries, archives, and museums (LAM)—while conveying models and approaches for collecting and preserving complex robotics material. The first phase of The Robotics Project (2019-2022), funded by the Alfred P. Sloan Foundation, aims to demonstrate that well-designed archival practices—incorporating interdisciplinary methods—can illuminate a theme as complex as a mode of scientific inquiry.

In this toolkit, we present the results of the first phase of our investigation. We introduce an approach to gathering crucial precustodial data from community members, stakeholders, and informants that supports future collection development, appraisal, accessioning, and subsequent preservation and description work. We apply this approach to robotics at Carnegie Mellon University, but we feel it has applications beyond robotics, especially those characterized by tangible products and collaborative processes, such as engineering and architecture.

By sharing this toolkit, we hope to cultivate a community of practice among archivists and information professionals engaging with robotics and similar fields.

MULTIMODAL ARCHIVIST

Today, archivists understand that our work demands often take us outside of traditional archival practice. This is especially true as we face the challenges of increasingly voluminous digital records and the complexity of hybrid collections that combine physical and digital material.

This toolkit uses the term ***multimodal*** to describe the interconnected ecosystem of tangible and intangible information, objects, digital artifacts, and narratives that comprise the scientific robotics research process. In our investigation of this ecosystem, we found a need to approach our work with a corresponding **multimodal mindset**, incorporating interdisciplinary methods and emphasizing preparation and data collection.

HOW TO USE THIS TOOLKIT

While library, archives, and museum professionals are recognizing the need to collect and preserve material from robotics history, few guidelines exist for these and other complex scientific collections. We hope to cultivate a **community of practice** among information professionals engaging with robotics and similar fields.

We invite you to seek out others preserving technology-driven fields to join an emerging community of practice. The information contained in this toolkit is designed for you to read, use, modify, and build on as we collectively develop new approaches to meet a growing need.

We aim to create a toolkit that will assist archivists and allied professionals in carrying out precustodial activities necessary for

- **collecting data that will inform collection development** and decisions concerning appraisal and acquisition
- **establishing relationships with communities** and potential donors
- **communicating archival practices** and missions to unfamiliar audiences
- **building the required interdisciplinary teams** to collect and steward complex hybrid collections

This toolkit will also aid anyone interested in

- **learning approaches for engaging with STEM** researchers, groups, and communities around archival topics
- **understanding audiences and use cases** for robotics material in archival, museum, and historical research settings
- **reviewing ethnographic methods for archival fieldwork**, including a 6-stage conversation guide
- **developing effective collecting strategies and policies** for robotics or similar fields
- **educating the next generation of archivists** and allied professionals in navigating the challenges posed by hybrid technology collections

2. APPROACH

BACKGROUND

In 2018, archivists completed a survey of our holdings in the Carnegie Mellon University Archives. Among the thousands of linear feet and terabytes of content, we were surprised to find a modest amount of material in our collections about the university's rich history in robotics research. Carnegie Mellon founded the first robotics department at a university—the Robotics Institute—in 1979, and since then, it has cultivated a global reputation in the field. We have a responsibility to preserve this material for its institutional significance, but it also has importance for research and study. To understand modern society means telling the story of robotics properly, akin to the scope, detail, and technical accuracy of historical projects that explore the development of atomic energy, the moon missions, or the earliest advancements in electronic computing. Historians, sociologists, anthropologists, archeologists, journalists, policymakers, and entrepreneurs are just a few of the user groups who will rely on robotics material to investigate the field's history, from its products and decisions to its scientists and communities of practice.

In recognition of this significance, we launched The Robotics Project, a collaboration between the University Libraries and the School of Computer Science, in 2019 to

- **assess robotics information artifacts and objects**
- **collect robotics archival material**
- **promote the material to broad audiences through research access and public programming**

Initially, the project team held informal conversations with members of the robotics community at Carnegie Mellon to let them know about the new initiative. These conversations revealed a large quantity of material in a wide range of formats—the software, code, video, photographs, websites, prototypes, and machines illustrated in **Figure 1**—with some material changing iteratively over time.

Artifacts	Research	Multimedia	Promotional
Machines	Proposals	Oral Histories	Announcements
Prototypes	Software	Audio	Newsletters
Parts	Code	Video	Events
Accessories	Data	Photographs	Conferences
Modules	Lab Notebooks	Scrapbooks	Proceedings
Tools	Documentation	Interviews	Presentations
Casings	Drawings	Graphics	Speeches

Communications
Reports
Policies
Correspondence and Emails
Course Materials
Websites

Figure 1. A summary of the material revealed in preliminary conversations with the robotics community, 2019.

For example, a video recorded on ¾" U-Matic tape may also exist in different versions on VHS, DVD, and as a digital file. We also found that material is hybrid, highly interconnected, and may only be understood in situ. The significance of hours of video footage documenting the build and structure of a robotic prototype may seem excessive, but in some periods of robotics research, video recordings and other visual documentation are vital to how roboticists execute, assess, and interpret the stages of their designs. During the research process, scientists and engineers frequently ***cannibalize*** prototypes and machines to pursue new iterations or entirely new models. A LIDAR sensor and chassis removed from an original self-driving vehicle may be repurposed, even decades later, to build a new unmanned ground vehicle. The remaining parts are then scrapped to make storage space for further innovations. In some cases, video is the only remaining visual documentation of the original.

Our team discovered it was nearly impossible to separate archival records from objects such as machines, computing devices, and prototypes more typically collected by museum-type organizations due to significant interconnections between these two categories of material. The provenance of material is also often difficult to discern, primarily due to the collaborative nature of robotics research and the considerable turnover and autonomy of individuals within labs. The landscape is further complicated by indeterminate copyrights, trade secrets, and other intellectual property concerns common to robotics research.

With this information, we began to assess our internal infrastructure to acquire and steward future robotics archival collections. At the time, our university employed a team of just three full-time archivists with little familiarity in robotics research. Our situation was further complicated by a significant collection backlog, overloaded facilities, and limited digital stewardship infrastructure.

In a procedural sense, we could theoretically review incoming material and deploy current best practices, but as we considered the copious amount of material, we began to identify immediate needs, such as defining our scope and creating an appraisal rubric aimed at answering a fundamental question: ***what is most important to collect?***

Our team determined we needed additional data to make informed collection development and stewardship decisions. We designed a two-part approach to address the precustodial steps we needed to take to gather information. We centered our efforts on two areas: **Interdisciplinarity and Community Engagement**. We review the insights afforded by this approach in **Section 3** and offer a Conversation Guide Model in **Section 4**.

INTERDISCIPLINARITY

The first component of our approach was to embrace an interdisciplinary team. As we explored the practice of robotics, we quickly recognized it as an interdisciplinary field. Robotics blends engineering, design, and computer science disciplines and contributes to medicine, agriculture, manufacturing, transportation, and many other areas. To effectively represent this expansive multidisciplinary enterprise, we needed to expand our team's expertise and identify peers in libraries, archives, and museums (LAM) tackling similar challenges. We also recognized the limited research available from LAMs on the collection and stewardship of robotics material and began seeking expertise outside of our peers and professional community. With support from the Alfred P. Sloan Foundation, we brought together a team with broad functional knowledge. We hired a processing archivist for robotics collections and sought colleagues in fields such as **history, information science, computer science, anthropology, and conservation**.

Social and Information Sciences

We had numerous conversations with colleagues in history, digital humanities, data management, archaeology, anthropology, user experience design, records management, software development, and cultural geography. Many of these experts were intrigued by the theme of robotics, but as we dove into the nuance of the robotics archive, they were eager to offer philosophical, functional, or methodological advice and insights. Working in a research library affords us ready access to this wide assortment of skills and interests. These conversations were invaluable to us during the formation stage and carried us onward as we began acquiring, processing, describing, and providing access to the material through the robotics archive at Carnegie Mellon.

Primarily, through these conversations, we developed a family of terms we could use to hone our approach, identify characteristics unique to robotics material, and establish a shared language among our team members. Inspired by data management, our team found the term ***ecosystem***, rather than the commonly used framing of a research lifecycle, better reflects the complexity and interconnectedness of individuals, communities, technologies, activities, information artifacts, and objects that comprise the research process. We also found ***stratigraphy***, a concept primarily used in geology and archaeology to account for context at an excavation site, helpful as we explored methods for capturing the in situ context of material in labs and other research locations. Lastly, the archaeological term ***chaîne opératoire*** or ***operational sequence***, which addresses the creation, use, and disposal of artifacts, helps identify and consider build processes and other robotics research activities.

Conservation

The stewardship of physical objects, particularly computing machines, presented a considerable challenge for us. While some members of our team had previous experience working with object collections in museums, our current operation was not well-suited for these artifacts.

Beyond the tangible needs of storage and gallery space, we encountered objects in varying poor conditions: damaged and dirty circuit boards, wires, batteries, sensors, gears, and wheels. We quickly generated many questions about conservation and similar areas.

- What is the ideal storage environment for these objects, and what methods are appropriate for transportation?
- What version of a machine should we acquire? What version has the highest historical value?
- Should we make repairs, replace existing parts, or maintain them in the condition and formation received?
- Is restoration a priority? Do we need the machines in working condition?
- Should the objects be displayed, demonstrated, or used as a teaching collection?

To help us with these questions, we hired a team of three conservators. This team was invaluable as they propelled our thinking on the care and stewardship of physical objects, addressing the challenges posed by electronic and mechanical parts and materials such as plastics and metals. They stressed that each machine requires a unique, customized plan and adequate resources. For example, if we aspire to document the development and progress of autonomous vehicles but do not have the conservation funds or storage space for large machines, we need to adjust.

They also raised philosophical questions that encouraged us to nuance and define our collecting strategy and intentions and consider future audiences. For example, if we intend to present interactive experiences in exhibitions, we must prioritize restoration in our conservation approach to get the robots in working condition. Alternatively, if we aim to showcase the development of the technology, we need to focus our attention on documenting the various iterations of machines over time.

The conservators urged us to consider documenting the intentions and decisions of the individuals and teams who create the machines. When acquiring a new piece, museums and galleries commonly document the ***Artist's Intention***, or the goals, designs, and plans of an artist concerning a specific artwork. This information supports future exhibitions and conservation work, especially for complex artworks such as technology-infused installation pieces and time-based media.[1] By documenting the ***Scientist's Intention*** as we talk with potential collection donors and assess robotics material for potential acquisitions, we gather crucial data that will assist with future processing, description, and access conditions.

These conversations confirmed our thinking that there are no one-size-fits-all solutions and that a clear collection development policy and intention are vital. The conversation guide detailed in **Section 3** integrates these insights.

Digital Curation and Preservation

Similar to the stewardship of physical artifacts, we needed to address knowledge gaps related to digital preservation and curation. While we have a long history of managing PDF and TIFF files, our team needed to consider integrating other formats, including data, code, software, video, and a wide variety of other born-digital content. We are not alone in these challenges. Many are widely known problems across the cultural heritage landscape, and we will grapple with them for years ahead.

[1] For information about best practices for time-based media and multimedia artworks see "Matters in Media Art," http://mattersinmediaart.org/.

To address our needs, we hired a digital preservation consultant with an overarching aim to share insights broadly across our digital infrastructure. Our situation involved scaling up our current capacities while also incorporating new formats. There is no single turn-key solution; instead, the task in front of us is to develop a constellation of solutions for various formats and configurations.

Obsolescence and version control are primary areas of concern. Technology develops and changes rapidly across formats and systems, complicated by legacy and modern third-party tools such as GitHub, Slack, Google Drive, and Overleaf. Furthermore, teams rather than individuals create this material. Proprietary platforms host project documentation, notes, messages, proposals, software, data, and drafts. Archivists found these records in filing cabinets in the past, but they exist on numerous platforms and with many individual users today.

Our challenge is also rooted in integration. Given the interconnected nature of robotics material, we wanted to design a holistic data model to reflect it. We sought to emphasize and demarcate the relationships between people, projects, and outputs. We set out to not only describe the material but to connect them, conveying the complex research ecosystem we were observing.

To evaluate the challenges ahead, we developed a prototype **Digital Robotics Archive**: https://roboticsarchive.library.cmu.edu/. We utilized Islandora 8 as a digital asset management system and web interface. The prototype forced us to find solutions and confront storage, display, and metadata questions. While we can look to theoretical frameworks such as the Open Archival Information System Reference Model (OAIS), it was not until we began building the prototype digital collection that the pragmatic and local infrastructure came to light. We recommend rapid prototyping if your project involves digital hosting, as it can allow models to evolve as you shape local policies and protocols.

Inclusive Collection Development

Themes of diversity, equity, and inclusion entered early into our consideration during the formation of The Robotics Project. Our team acknowledged that inclusivity would be critical to formulating future robotics archival collections: from our appraisal efforts and what we accepted into the collection to how we process, describe, and provide access to the material. In this spirit, we affirm that archives and archivists are not neutral and that archival appraisal and collecting practices historically have favored white, male, and administrative/institutional perspectives. We acknowledge robotics itself has a reputation for being a white, male-dominated field, but there are many more stories to tell. We are committed to telling robotics stories highlighting the experiences of under-engaged and marginalized groups.

The research enterprise at Carnegie Mellon is a complex, ever-changing ecosystem. Robotics labs comprise undergraduate and graduate students, postdocs, staff, lab managers, interns, visiting researchers, research faculty, and teaching faculty. This community operates against the backdrop of a broader network of administrators, grant managers, business officers, communications experts, compliance agents, contract officers, Institutional Review Board members, IT managers, consultants, manufacturers, companies, funders, government agencies, and users, helping to propel an ongoing stream of initiatives.

In a university setting, typically the principal investigator (PI) of a funded project or the head of a lab is the most public member of a team. To document robotics research accurately, we need to collect beyond the PI and engage the students, engineers, administrators, and staff who build machines, write code, purchase equipment, maintain facilities, and make the work possible. Our collecting efforts should reflect the backgrounds and contributions of the research team members.

We enlisted the help of a senior archives consultant to explore inclusivity and community engagement for collection development. We sought to design an approach, outlined in **Section 2**, that would help us build relationships with the robotics community, engage with the many perspectives and experiences of robotics research, and gather crucial insight into robotics material. We outline the findings of this approach in **Section 3**.

A Note on Team Science

We would not have been able to tackle the first phase of our research in any meaningful manner without a fusion of ideas and methods from other domains. As archives, libraries, museums, and other cultural heritage organizations aim to build, manage, and sustain collections for modern audiences, we need to consider and adopt modern knowledge frameworks.

We have the benefit of learning from other professionals further along on this journey, namely in the health sciences. We recommend *Collaboration and Team Science: A Field Guide,* sponsored by the National Institutes of Health. This booklet outlines key aspects beneficial to any team, and it serves as an approachable primer before embarking on a new program or project.

Some of the key aspects the authors recommend include **developing trust, creating a shared vision, outlining clear roles and responsibilities, ideally with a written agreement determining recognition and credit, and understanding that conflict will arise and therefore determining a decision-making protocol before difficulties emerge.**[2]

---

[2] Bennett, Michelle L., et al. *Collaboration and team science: a field guide*. National Institutes of Health. https://www.cancer.gov/about-nci/organization/crs/research-initiatives/team-science-field-guide/collaboration-team-science-guide.pdf Accessed on September 25, 2022.

The most important and tangible advice we can offer is to cultivate a **growth mindset**. Phase 1 of The Robotics Project expanded our thinking, and we recognize that future archival work, particularly in STEM fields, requires more interaction and integration with allies, partners, audiences, and communities as a collaborative process. This first step is creating bridges with colleagues working across digital scholarship, data management, publishing, scholarly communications, and open science. The Robotics Project has inspired us to advocate for and incorporate deeper interdisciplinary practices in our wider professional endeavors.

COMMUNITY ENGAGEMENT

In addition to consulting a wide range of experts to bolster our functional knowledge, we sought to build our understanding of robotics through community engagement. We incorporated insights from our interdisciplinary team to design a community engagement approach that would use ethnographic methods to

- **help us understand the products and by-products created by robotics**
- **gain insights into the vast community of people who carry out and participate in the research process**
- **build strong relationships with potential donors, stakeholders, and informants**
- **partner with the robotics community to develop a shared vision for future outcomes of The Robotics Project**

Our approach centers on *precustodial fieldwork* by archivists. We define precustodial (also sometimes written pre-custodial) as the phase of activities before acquiring and accessioning material into an archive or other repository.

An archivist engages in precustodial activities before taking custody of any material (or applying a shared or post-custodial model). This work can include site visits, record surveys, meetings with prospective collection donors and other stakeholders, and recorded interviews with prospective collection donors and community members. In some cases, it may also include consultations with experts in other fields and public history activities such as educational programs, exhibitions, events, workshops, seminars, and even podcasts, if the activities engage communities and individuals around collection development. All of these activities may or may not lead to future acquisitions, but they are valuable tools for building relationships, increasing the reputation of the repository, and assessing available material for collection development.

Published guidelines on ***precustodial fieldwork*** are largely missing from modern archival literature.[3] We can attribute this at least in part to the relative lack of emphasis on this work in the archives profession. Our primary responsibility is custodial, and not all archivists are in a position to conduct fieldwork. Sometimes this work is reserved for specialized positions or high-level administrators in an

[3] Yun, Audra Eagle and Society of American Archivists, *Archival Accessioning*. (Chicago: Society of American Archivists, 2021), 35.

organization. Many of us, however, including those working in smaller organizations and repositories, will find ourselves engaging in this work out of necessity. Precustodial fieldwork has the potential to help archivists fill gaps in our existing collections and gather data that will aid in future accessioning and processing activities (i.e., the process of acquiring and cataloging archival material). Likewise, larger collecting organizations may benefit from engaging in fieldwork, such as the process proposed by this guide, after performing collection assessments to determine new collecting areas, especially as they start to contact prospective donors and perform record surveys.[4] As always, archivists must weigh the benefits of engaging in precustodial fieldwork with their regular custodial activities (i.e., preserving and providing access to current collections), as well as any limitations on their available resources, such as time to perform the work and space for new collections.

Our recommendations stop short of ***precustodial intervention***, an idea Australian archivist Adrian Cunningham proposed in 1994. Recognizing the rapid rise of digital records and the need for change in the profession, Cunnigham suggested archivists collecting personal papers approach donors early in their careers and take an active role in designing "recordkeeping systems" to promote standardization and make sure donors are creating adequate documentation of their work.[5] He concedes that there are challenges to this approach, namely getting donors to agree to a lengthy process, but he "laments [he] can see no other alternative if personal records archivists are not to be made redundant antiquarians by the relentless march of technology."[6] Thirty years later, archivists are still facing these challenges, but

---

[4] Weideman, Christine, "A New Map for Field Work: Impact of Collections Analysis on the Bentley Historical Library." *The American Archivist* 54, no. 1 (1991): 55.

[5] Cunningham, Adrian, "The Archival Management of Personal Records in Electronic Form: Some Suggestions," *Archives and Manuscripts* 22 (May 1994): 100.

[6] Cunningham, "The Archival Management of Personal Records in Electronic Form," 101.

our team does not have the resources to pursue donor relationships at this level of granularity. We believe many archivists are in similar situations, and more research is needed to identify a practical process for "precustodial intervention." Instead, we propose an approach to precustodial fieldwork that allows archivists to gather data about potential archival material without intervening in its creation or maintenance.

That said, the recommendations outlined in this toolkit naturally lend themselves to **proactive collecting**, in which an archivist actively seeks out new collections and relationships with informants and potential donors. It could also apply, however, to **reactive collecting** situations where an archivist is gathering information about a donor or collection arriving at the repository voluntarily. In these cases, archivists must weigh the benefits of undertaking this work with the urgency of the situation in front of them.

Pulling together insights from our interdisciplinary team and consultants, we designed the fieldwork method detailed below. See **Section 3** for the findings of the ethnographic study.

Ethnographic Study Methodology

Our work began with numerous months spent building rapport, becoming familiar with the relevant administrative, policy, and technical contexts, and conducting unstructured, informal conversations with potential donors, stakeholders, and informants. This preliminary work enabled the design of a robust **methodological data collection and analysis strategy**.[7] At the end of this initial period, we

[7] Yin, Robert K, *Qualitative research from start to finish* (Guilford publications, 2015), 306.

finalized our data collection and analysis strategy, secured approval for human subjects research from our university's Institutional Review Board, and began our research effort.

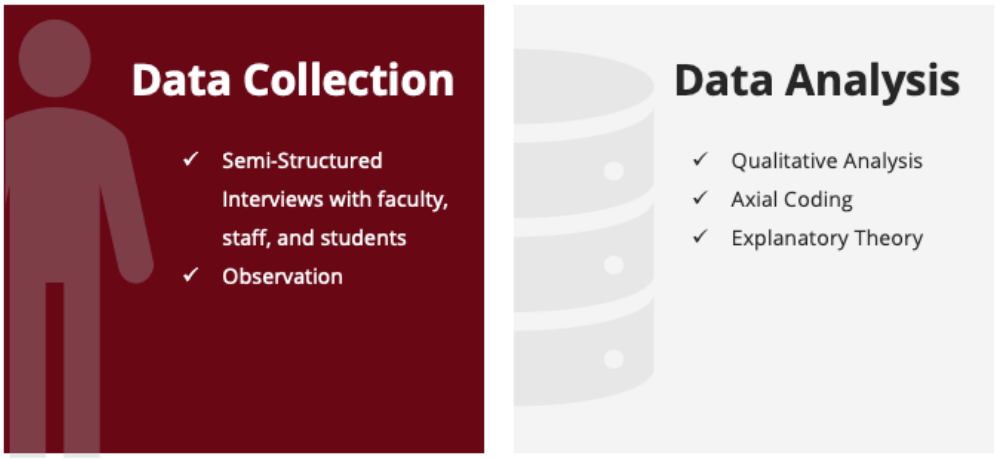

Figure 2. A summary of our field study methodology.

We designed the study to unfold in three parts (See **Figure 2**). First, we collected ethnographic and semi-structured interview data. We then used a grounded theory analytical approach, which aims to develop new hypotheses based on the data gathered, to produce a holistic account of the archival products and byproducts of robotics labs at Carnegie Mellon, including

- **comprehensive account of archival products** and byproducts
- **contexts where these are accessed** (who, what, when, and where)
- **products and byproducts at risk of loss** and degradation due to factors such as physical location and digital obsolescence
- **context about archival value** (e.g., historical, evidential, informational)
- **context on the communities** (who, what, when, where, and why) and collaborative research processes in the robotics ecosystem at Carnegie Mellon University

Finally, we used the data to develop this toolkit encompassing archival practices for use by curators, archivists, librarians, or others who may steward scientific and technology collections in libraries, archives, and museums.

Data Collection

In this study, we selected a convenience sample of three robotics labs at Carnegie Mellon. A convenience sample of high-ranking representatives from each lab was invited to participate in the study via email. Then, we used snowball sampling to identify and recruit at least three lab members to participate in data collection efforts. We strove to recruit participants from each lab to reflect the diversity of roles involved in robotics research at the university, including faculty, staff, and students, to capture diverse perspectives on the research process and the resulting archival products. Participants were asked to commit approximately one to three hours to the study: no more than 10 minutes for an informed consent process that our institutional review board had vetted, no more than 60 minutes for participation in a semi-structured interview, and no more than 60 minutes to engage in an observation session. All participants were current employees or students of Carnegie Mellon University involved in the robotics research process at the time of the study.

Semi-structured interviews (See **Appendix** for the interview script) with members of each of the labs were collected and recorded using videoconferencing technology and transcribed using automated voice-to-text transcription software. Semi-structured interviews with interactive prompts allow the researcher to gather information directly from those engaged in relevant activities, while offering flexibility to pursue new ideas that emerge in the course of an interview.

Interviewers remained open and flexible to pursue important ideas that arose during the interview process but were not contained in the prompts. Participants were interviewed about their experience of the robotics research process and the resulting archival products and byproducts of their labs. Semi-structured interviews are a key element of our grounded theory approach. ***Grounded theory*** is a systematic, iterative approach to theory-building in which insights are derived from an open and responsive approach to data collection and analysis, which allows participants to drive discussion and organic exploration of emerging topics.[8]

In-person ethnographic **observation** sessions were also arranged with each of the labs. Ethnographic methods allow the researchers to paint a realistic and detailed picture of the context, actors, activities, and artifacts in which robotics research unfolds.[9] The purpose of these visits was to observe the working environments of the labs and gather data on the products and byproducts in the physical spaces. During each ethnographic observation session, we used an established heuristic framework for collecting relevant observational data, including contextual details of **activities, environments, interactions, objects, and individuals.**[10] The field notes resulting from observational sessions were used to inform and extend the analysis of semi-structured interview data.

---

[8] Strauss, Anselm L and Juliet M Corbin, *Basics of Qualitative Research: Grounded Theory Procedures and Techniques* (Newbury Park, CA: Sage, 1990), 317.

[9] Yin, *Qualitative Research From Start to Finish*, 150.

[10] Hanington, Bruce, and Bella Martin, *Universal methods of design expanded and revised: 125 Ways to research complex problems, develop innovative ideas, and design effective solution*s (Rockport Publishers, 2019), 14.

Data Analysis

We employed a grounded theory approach to the analysis of semi-structured interview transcripts and ethnographic field notes. We began by applying a modified framework, identifying segments of text as associated with **products, contexts, ecosystem, sustainability, obsolescence, value, connections, function, culture, works, and improvement.**

Then, we used an ***axial coding*** approach by reviewing all text segments associated with one of these dimensions to identify dominant themes in our interviews.[11] This approach enables iterative development of ***explanatory theory*** and has been used previously in library contexts.[12] This approach resulted in a descriptive account of the archival products and byproducts of robotics labs.

---

[11] Vollstedt, Maike, and Sebastian Rezat, "An introduction to grounded theory with a special focus on axial coding and the coding paradigm," *Compendium for early career researchers in mathematics education* 13 (2019): 81-100.

[12] Harati, Hadi, et al. "Factors affecting the unplanned use behavior of academic libraries users: Towards an axial coding pattern." *Aslib Journal of Information Management* (2019).

3. FINDINGS

INTRODUCTION

In **Section 2**, we presented our two-part approach—interdisciplinarity and community engagement—and the methodology used in our ethnographic study of labs at Carnegie Mellon University. We gathered necessary insight into the robotics research process, examining the actors, products, and processes relevant to answering the following questions:

- **What are the archival products and by-products (e.g., records, materials, artifacts) of robotics labs at Carnegie Mellon?**
- **How are these products created? Who created them?**
- **What formats and conditions are they in?**

We sought to use the data gathered in the ethnographic study to inform appraisal and collection development decisions deployed at Carnegie Mellon University.

In **Section 3**, we share our findings. Our work to study the data is ongoing, but we completed an initial qualitative analysis in 2022. Although our analysis does not tell us the prevalence of these results within the robotics research ecosystem as a whole, it reveals their presence within our organization's context.

We are sharing our findings so that others can use these insights as a jumping-off point. We aim to supplement existing literature and guidelines for archivists on documenting and preserving science and technology fields.

We outline our findings in the following three parts: **Products and By-products of Robotics, Prospective Functions and Audiences, and Social Networks.**

PRODUCTS AND BY-PRODUCTS OF ROBOTICS

Our initial conversations at the university revealed large quantities of material in various formats (See **Figure 2**), and this study clarified additional details about the products and by-products of robotics research. Some field-specific knowledge about these products is crucial for archivists determining collecting priorities and appraisal strategies.

Unsurprisingly, participants described **datasets, journal publications, and conference proceedings** as high-value to the field and central to their success. Several participants indicated that these products contain the necessary information to preserve and communicate their work.

All labs in this study had machines, robots, and prototypes fundamental to their research. Most participants indicated that physical products such as these, as well as supplies and tools, were used

communally by all lab members. One participant said their lab inherited their robot from another, explaining that it is expensive and requires time to develop knowledge to run, so they try to reuse or repurpose it whenever possible. Several participants expressed that it is difficult to dispose of physical machines and parts once they have become obsolete.

All labs in this study maintained some central repositories that could be accessed and used by lab members, but the forms and functions of these repositories vary widely. One participant described a master spreadsheet used to track all significant lab research activities, while others said GitHub was their primary repository for documentation and code. YouTube, Google Drive, and DropBox were also mentioned numerous times. Participants indicated they use these repositories to store videos and photographs, datasets, code, and other information such as lab supply inventories, team member lists, and procedures. Several participants explained they use wikis on various platforms to supplement their documentation and share knowledge among lab members.

Other secondary forms of documentation were crucial to communication and task management among team members and collaborators. Participants gave examples such as Slack, Overleaf, and Confluence but indicated these communications vary widely from lab to lab and from project to project.

Crucially and of interest to archivists, **sustainability** is a common theme throughout the interviews. Nearly all participants expressed concerns about long-term reproducibility and preservation due to obstacles in maintaining project documentation and knowledge among lab members. Teams change annually or every semester as students and post-doctoral research assistants graduate or move to other positions. This creates significant challenges for scientists and archivists alike concerning lab data and documentation.

PROSPECTIVE FUNCTIONS AND AUDIENCES

Functions

For decades, archivists have used functional analysis and macroappraisal to determine collecting priorities by examining the activities undertaken by individuals within an organization. Perhaps one of the most enduring examples comes from Helen Willa Samuels's functional study of universities, *Varsity Letters: Documenting Modern Colleges and Universities.* In this study, published in 1992, she proposed seven functions of universities: **Confer Credentials, Convey Knowledge, Foster Socialization, Conduct Research, Sustain the Institution, Provide Public Service, and Promote Culture**.[13] These categories have guided generations of university archivists in adequately documenting their institutions.

Although functions are not the only rubric used by archivists to inform appraisal and collecting decisions, it continues as part of the appraisal toolkit, often appearing alongside factors such as user interest and barriers to use.[14] For our purposes, we sought to understand how the functional analysis approach used by Samuels and her contemporaries might be altered by modern, technology-driven ecosystems like those of the robotics labs at Carnegie Mellon.

In our study, participants describe various activities directed toward the functions identified by Samuels, primarily Convey Knowledge, Foster Socialization, and Conduct Research. Faculty participants indicate activities involved in conducting research, educating students, managing

---

[13] Samuels, HW, *Varsity Letters: Documenting Modern Colleges and Universities* (Lanham Md. and Chicago Ill: Scarecrow Press and The Society of American Archivists, 1998), 22.

[14] Yun, *Archival Accessioning*, 41.

resources, and providing public services such as advising on public policy. Likewise, student participants report activities such as publishing papers, conducting site tests, and managing labs. Staff participants describe similar activities but are more focused on activities sustaining their individual labs, such as building, programming, and deploying robots.

There is one area, however, that is not accounted for in Samuels's design: **entrepreneurship**. Participants from all three groups—faculty, staff, and students—describe an environment ripe with entrepreneurial activity and extensive interactions with private companies. One participant expressed that it sometimes feels like a business inside of a university when discussing the culture and environment. Another participant explained that faculty can work at a private company for a certain percentage of their time, while others described starting companies of their own to commercialize research they began at the university. One participant explained further that they sought to create private companies to keep student talent in the region rather than lose it to other locales. Given its prevalence in the data, our team will incorporate entrepreneurship activity in our appraisal and assessment of collection material.

For modern research universities, especially those engaged in scientific research, we recommend archivists consider the function *Employ Entrepreneurship* as you assess your institution if you are using approaches such as functional analysis or macroappraisal.

Audiences

Archivists consider future researchers and users when appraising potential material. In our study, faculty, staff, and student participants struggled to identify the value of our robotics collecting initiative. Although all participants supported this work, they found it challenging to articulate positive outcomes. Despite this, the data reveals several prospective audiences or user groups for robotics archival material. Archivists will notice many traditional users of archives missing from this list, such as social scientists and the media.

We grouped the audiences identified through this study into the following six categories: **Nostalgia, Marketing, Research and Administration, Education, Public, and Libraries, Archives, and Museums**.

Nostalgia

There is a desire in the robotics community to share stories about their past. As one participant put it, there are so many interesting stories. Another participant explained they were looking for validation of their belief in the value of robotics history, while later describing their desire for current generations to understand how difficult robotics research was compared to today.

Marketing

Numerous participants identify significant marketing potential. Participants referenced the renown of peer institutions and noted a desire to build the public identity of robotics at Carnegie Mellon University. One participant explained that the public does not talk about robotics research happening at the university, pointing out specifically that the university should share more about its history.

Research and Administration

Although most participants struggled to identify relevance to their work, others discussed how archival material would be helpful in their research and administrative processes. Several participants discussed the negative impacts of turnover and the lack of accessible documentation on critical areas such as institutional memory, lab management, and the transferability of research to new students. Others explained that access to archival robotics material would help them complete literature reviews, a necessary first step in the research process, and track robotics research trends over time, helping them identify new or underserved research areas. Lastly, one participant sees potential for robotics archival material to inspire administrative changes at the university and improve conditions for students, while another wanted assistance clearing decommissioned robots from a hallway.

Education

Several participants discussed challenges related to current robotics students. Although participants offered no explicit examples of the pedagogical application of archival material to robotics education, some participants expressed the need for students to understand more about the history of robotics and the development of the field, especially given its rapid growth. In one instance, a participant noted that access to archival material and robotics history would help students find direction and inspiration in their own research.

Public

While one participant shared that they do not think the public will be interested in archival robotics material, most participants expressed the need for greater public engagement and sharing. Two participants want to see a museum for robotics history, while others would like to see a publicly accessible database.

Libraries, Archives, and Museums

Participants acknowledged the need for experts in libraries, archives, museums, and similar professions to help them preserve robotics history. One participant pointed out that they could not preserve material alone, while another indicated they needed experts to help determine what material has long-term value.

Although these audience categories are not comprehensive, our team will consider them as we develop collecting policies and make appraisal decisions. They provide insight into the value of archival work for the robotics community.

SOCIAL NETWORKS

Our study provides insight into the impact of social networks on the robotics research process. Participants consistently discussed the power and importance of collaboration within the university and across institutions. One participant noted that a new research area resulted from a previous friendship and serendipitous, casual conversation, while another remarked on available grants from Carnegie Mellon specifically designed to incorporate researchers from multiple institutions. That participant went on to explain that some collaborators can lend needed credibility to research projects. Awareness of these social networks is essential for archivists looking to collect robotics research material.

It is important to recognize the impact this collaborative environment has on the distribution and availability of future archival material.

Archivists collecting this material may find geographical distribution across buildings, campuses, cities, and countries. In our experience, a single research project may involve, for example, five universities in three states and two countries. It is not practical to accession material from all of these institutions—due to the widely varying retention schedules and legal mandates across organizations—but we can note this distribution in the finding aid. It is a simple but effective step we can take toward documenting the collaborative environment.

4.

THE ROBOTS CONVERSATION GUIDE FOR PRECUSTODIAL FIELDWORK

INTRODUCTION

The ROBOTS Conversation Guide provides an easy-to-remember acronym describing six stages of relationship development in precustodial archival fieldwork: **Research, Outreach, Build, Offer, Transfer, and Steward.** Resulting from our two-part approach described in **Section 2** and our findings outlined in **Section 3**, ROBOTS aims to create guidelines for archivists engaged in collecting multimodal material that requires significant preparation and investigation prior to acquisition. Although this guide will not apply to all precustodial fieldwork undertaken by archivists, we aim to contribute to a growing need for guidelines in this critical area and cultivate a community of practice among archivists collecting in robotics and similarly complex fields such as computer science, engineering, and architecture.

We designed the guide with maximum flexibility in mind. The stages may occur in a single meeting or evolve over many years. The stages also take into account the steps that an archivist must consider when acquiring a new collection, such as securing legal agreements that hand over ownership of the

material and physically or digitally transferring the material to the archive or repository, but this guide is not comprehensive. It should not replace the acquisition policies or procedures of your organization.

This guide is modeled in part after **oral history methods** for conducting interviews intended to capture personal commentary on historical events. Oral historians follow a set of principles and best practices, such as those recommended by the Oral History Association, aimed at maintaining respect and transparency between all participants.[15] The oral history process typically includes three steps: **prepare, interview, and process**. At their core, these steps involve the careful gathering of research on an individual or topic, crafting that information into a thoughtful set of questions or topics, recording the interview conversation, and creating transcripts or other documentation that are deposited along with the recording in a library/archives or shared publicly in some other way. In oral history practice, interview questions are typically open-ended and intended to avoid simple yes or no answers. Interviewers avoid jargon, idioms (e.g., keeping the lights on, hitting a home run), and other phrases that assume a shared cultural experience or vocabulary. By employing these and similar methods in this guide, we aim to arrive at an ethical and productive experience for all involved parties. This guide also borrows from other qualitative research methods, such as ethnographic observation and thematic analysis. For more information about our ethnographic study method, see **Section 2**.

[15] "Principles for Oral History and Best Practices for Oral History," Oral History Association, adopted October 2018, https://oralhistory.org/principles-and-best-practices-revised-2018/.

The ROBOTS Conversation Guide contributes to a growing need for literature that addresses precustodial archival fieldwork. It presents a set of practical stages and questions—inspired by ethnographic methods—that we can apply as we approach donors, stakeholders, and communities about collection donations and other archival activities.

CONVERSATION STAGES

Stage 1: Research

Precustodial fieldwork starts with research. Familiarize yourself with as much information about your topic as possible. Perhaps you arrived at the topic through an assessment of the collections in your repository that showed a gap in your current holdings, or maybe you received a referral from someone inside or outside your organization. Regardless, in this first stage, it is crucial to discover as much as possible about the topic you want to explore. Consult a variety of resources, and examine your repository's holdings for related material. This initial research will help you identify potential donors, stakeholders, and informants. It will also help you avoid collections that may not be an appropriate fit for your repository, and once you start initiating conversations with potential donors, build rapport with them.

During your research, it is equally important to examine your available resources and determine if it is within your scope. Consult your administrators and colleagues for information about the current capacity of your repository, and ask them if they have had contact with any potential donors related to the topic. Also, talk with your administrators about available resources to support your fieldwork.

Sample Questions

At this stage, you can ask yourself and your colleagues the following questions to assess the topic you have identified.

- How does this align with our current collecting strategy or policy?
- How will this impact available resources such as facilities, funding, and personnel (including myself)?
- Do we have the infrastructure to support new collections resulting from this fieldwork?
- What costs can we predict will come from conversations with potential donors?
- Can I reach out to individuals or communities related to this topic, or would I benefit from introductions?

Stage 2: Outreach

In Stage 2, start reaching out to the individual, organization, group, or community. If your initial contact with them is via email or other written communication, aim to arrange an in-person meeting if at all possible. If that is not available, which may be the case due to distance, time, or health, arrange a phone call or video call. If possible, do not conduct all conversations via email. Ideally, arrange meetings with individuals, but group conversations could also be helpful, depending on the

circumstances. For example, talking with a couple or a team may provide additional information as they may help each other remember specific details.

This initial conversation is exploratory. Based on your research, come to the initial meeting prepared with topics or questions you want to address. You can create a formal list of questions or take a few notes. Regardless, it is important to plan in advance how you would like the conversation to go. Also, come to the meeting prepared to answer questions about your background and work. If you are just starting with fieldwork, you may want to create a short **elevator pitch** that explains your work to a broad audience, such as those recommended by the Society of American Archivists.[16]

Before beginning a conversation, it can be helpful to examine your own assumptions and be prepared for surprises: we may learn things in the course of conversation that open whole new avenues of inquiry or overturn assumptions. We are better equipped to recognize these insights if we take a moment to prepare ourselves beforehand.

Avoid using jargon, acronyms, idioms, and phrases that assume shared cultural backgrounds and experiences in your conversation. Start with open-ended questions that avoid simple yes or no answers, and follow up with more specific questions based on their response. Practice active listening and be respectful by avoiding interruptions.

It is essential to take notes during the conversation. You may want to ask a colleague to attend the meeting with you and take notes so that you can give your full attention to the conversation. Write

---

[16] "Crafting Your Elevator Speech," Society of American Archivists, updated December 2017, https://www2.archivists.org/advocacy/public-awareness/elevator-speech.

down names, themes, and recurring topics, such as the names of their collaborators and team members, including administrative and engineering staff, and the titles of relevant articles or other resources. Pay special attention to their language, including **terminology, jargon, acronyms, and phrases**, and note how these are related to the topic. You can also ask them to explain terms that you do not understand. Sometimes interviewees do not recognize when they are glossing over a topic that may play a key role in the conversation.

After the conversation, review your notes and write down any additional thoughts. You may also want to create a **prospective donor or lead file** containing your research, notes, and other relevant material. In the future, if your fieldwork results in a new collection donation, this file will form the basis for the initial accession record and inform future processing activities. For more information about common methodologies and practicalities of collecting qualitative data, including note taking, we recommend literature such as *Qualitative Research From Start to Finish* by Robert K. Yin.

Sample Questions

In Stage 2, ask questions about their general background and experience, such as

- Tell me about your background.
- When did you start working with [topic]?
- Can you tell me about the work of [topic]?
- Is there anything we should read to learn more about [topic]?
- What accomplishments are you most proud of related to [topic]?
- What do you think is most important to preserve related to [topic]?
- Who did you work with on [topic]?
- Is there anyone else we should talk with about [topic]?
- When can we talk again about [topic]?

Stage 3: Build

Stage 3 builds on the information gathered in Stage 2. Continue employing the same strategies used in the initial conversations, such as creating a list of topics or questions for each meeting, asking open-ended questions, avoiding jargon and confusing language, and taking copious notes. As you build up the relationship, you may gather information quickly in just a few conversations, or it may take place over several years.

This stage introduces questions about potential collection material. Using the information you have learned about the topic, ask questions about format, quantity, and other details that will help you determine if the collection is within your scope. You can also assess the material's historical or other value and condition and forecast future expenses or hurdles such as **transportation, conservation, storage, and complex copyrights**. If possible, ask for a rough inventory or summary of the material.

This stage can be difficult due to differences in **communication**. For example, when doing fieldwork for our project, our team was challenged by language differences between roboticists and archivists. Archivists and roboticists use typical archival terms such as documentation, records, and archive differently, causing repeated confusion in conversations. It is essential not to take the definition of the words we use for granted. In your conversations, pay attention to the terminology used by the community and aim to arrive at a shared language to talk about archival work and preservation. This phenomenon has been explored deeply in History of Philosophy of Science literature, such as *Structure of Scientific Revolutions* by Thomas S. Kuhn.[17]

---

[17] Kuhn, Thomas S and Ian Hacking, *The Structure of Scientific Revolutions* (Chicago: University of Chicago Press, 2012).

During this stage, you may also consider asking for a guided tour of the room, studio, lab, or workspace where the work happens or the materials are stored. For example, our team arranged site visits for several robotics labs where we explored the environment and the creation process. We were also able to assess material in person and ask more detailed questions based on our observations. You may want to ask to create an audio or video recording of the tour.

As you build the relationship, use these conversations as an opportunity to continue talking about archives and your work. Share information about the archives profession and the current collections in your repository. It will help potential donors build an understanding of archives and set appropriate expectations if the conversation results in a collection donation. Explain the differences between archives, libraries, museums, and other collecting organizations if needed.

Continue taking notes throughout this stage and add this data to your potential donor file. If you arrange a site visit, you can also add any recordings and notes to the potential donor file.

Sample Questions

In Stage 3, ask the individual for information about specific projects to better understand the creation process and any potential resulting material. You can also ask questions about the material's value, quantity, and format. These questions could include

- Can you tell me about one significant project or product of your work?
    - What was the build/creation process like for that project or product?
    - What or who was the impetus for starting it?
    - Who else was involved?
    - How did you communicate with each other? (e.g., notes, emails, reports)?
    - Do you have any photographs or videos related to it?

    - Where did the work take place?
    - Who was your intended audience?
    - What hardware and software were used or written to develop it?
    - What or who defined the endpoint of the project?
    - What happened after the identified endpoint?
- Do you have anything from your early years working on [topic]? Or the early years of this work?
- How much material do you have? Can you share an inventory, description, or rough list?
- What would you like to see come out of our efforts to preserve material related to [topic]?
- Is there anything else I should have asked you about that I did not?

Stage 4: Offer

In Stage 4, ask if they would consider donating the material. This stage aims to secure the donation and **manage expectations** regarding future preservation and access to the collection. When asking for the donation, explain your interest in the material and why it is a good fit for your collections. You can also reference specific items you discussed with them previously.

It is critical to review the donation process and any requirements of your organization and ensure that they align with the donor's expectations. If the collection requires a deed of gift or other agreement, thoroughly review the process with them. Discuss intellectual property rights and any access or use restrictions the donor may want or need to place on the material. Depending on your organization's policies and procedures, involving colleagues from other departments, such as legal, may also be appropriate. Consider contacting development or advancement, as the donor may be willing to consider a financial gift to support the collection. You can explain the costs associated with archival work and long-term stewardship to the donor if appropriate.

You can also begin to assess transportation or transfer costs for the material. It will depend on location, format, quantity, and other factors. It is crucial to get a clear picture of the collection and its current location at this stage.

Sample Questions

In Stage 4, ask questions concerning the donation of material.

- Would you consider donating [the material] to our preservation efforts?
- Who owns [the material]?
- Where is [the material] currently located?
- What kind of condition is [the material] currently in?
- Is [the material] at risk of damage during transfer?
- How would you like to see [the material] transported?
- Is there anyone else we should talk to about [the material]?
- Do you have any questions about the [deed of gift] process?

Stage 5: Transfer

In Stage 5, transfer the material to your facility and take custody of the collection. Make the arrangements with the donor, and determine who will fund transportation. Many donors assume the archives or repository will cover transportation fees, so it is important to establish shared expectations.

It is also important to begin gathering detailed information about the condition of the material. You may consider creating an **acquisition form**, especially if the collection involves multiple modalities. For example, as our team worked to collect material from the history of robotics at our organization,

we found it necessary to gather detailed information about related mechanical components and digital dependences because these factors impact our preservation efforts. You can send the form to the donor to fill out, or depending on the circumstances, fill it out yourself throughout your conversations. Add the form to the donor file.

Sample Questions

During this stage, ask the donor questions about the best mode for transfer. You can also create an acquisition form containing questions about the context and technical details of the material. We developed the questions below in collaboration with our team of conservators, so they address the specific needs of robotics objects, but they could also apply to other hybrid and scientific collections.

- Provide a brief narrative description of what physical functions the robot performs (e.g., what does an observer see). Include details that describe the robot's appearance, movements, and aural or olfactory elements (intended or not).
- Is there media available that demonstrates the object functioning? (What format(s) are available, and can they be made accessible to researchers and/or included in the archive?)
- Who was the defined end-user or audience of the robot? (e.g., civilian, commercial, military, etc.)
- Is the robot, or another iteration of the robot, being used today? How and by whom?
- Did/does the robot function precisely as the creator(s) intended? In what ways did/does its actual function differ from its intended function?
- Provide a technical description of how the robot functions from the point of activation (i.e., turning it on, start-up, etc.) to completion (i.e., end of function, powering down, etc.). How is the robot operated (i.e., how does it turn on/off)?
- What additional components (i.e., computers, remote controls, human, etc.) apart from the robot itself are required for it to perform its function?

- In addition to the physical functions, use this space to describe or diagram the technical constituents and their relationships with one another (or refer to a document that does).
- Is the robot intended to operate within or for a predetermined duration? How was/is this duration defined?
- If no set duration, what is the speed at which the robot is intended to perform its function(s)? What methods were used to determine and set its pace, and how was this measured?
- Are there different sequences/programs that the robot follows? If yes, please describe and explain how a sequence is selected or performed.

Stage 6: Steward

In this final stage, shift your focus to stewarding the relationship with the donor, group, or community. The stewardship stage does not end until the donor's relationship with your organization ends. Sometimes this can lead to formidable **social overhead**. Ongoing activities might include sending ongoing updates to the donor, hosting events and private tours, and working on rights management and future accruals. Every collection requires meetings, emails, phone calls, and other streams of contact with donors and stakeholders that will take time on your schedule. Engaging advancement, development, or other colleagues in your organization may be helpful in managing this stage effectively. It is important to consider the hidden costs of social overhead as part of long-term stewardship.

5. CLOSING CONSIDERATIONS

Robotics presents significant challenges for archivists and information professionals. During the course of our work, we encountered many questions that we anticipated, such as handling digital records, rights management issues, and representation. As we reflect on the multi-year journey to develop The Robotics Project, we want to offer some closing considerations for others exploring multimodal archives.

Is it a Project or a Program?

Our work in robotics began with the recognition that we had gaps in a particular area that is significant at our university—robotics. Similarly, you may identify subject areas that you want to address. When approaching multimodal disciplines such as computer science, engineering, or similar, it can be challenging to collect material at scale. **The critical question to ask first is: *what are you building?***

We recommend spending ample time on this question. Before you think about workflows, policies, digital storage, or a collecting strategy, you need a clear vision. Are you looking to grow the archive by selectively acquiring new donations, or are you building a thematic collection? In the first scenario, you are looking for a range of records to fill gaps in your existing holdings. In the other circumstance, you are launching a notable initiative requiring ample time and resources. Understanding what you are building will inform the level of investment needed.

As we deepened our understanding of robotics, we recognized that the value was not in the records alone but in the connections between the elements that make up the robotics ecosystem—**the teams, individuals, and outputs of the research process**. We discovered The Robotics Project would require more than just additional space for boxes and bytes, which had a ripple effect beyond the archives to the rest of our organization.

If you decide to prioritize multimodal collections, it will likely impact your organization's infrastructure. Help your colleagues anticipate the impact by having a compelling and united ethos and collection development intention. We found these factors crucial to achieving our goals.

Pragmatic Partnerships

Community engagement (See **Section 2**) is critical to developing and sustaining multimodal archives. This approach enabled us to observe, understand, and appreciate how scientific knowledge is shaped and created. When facing vast quantities of records, this insight helped us determine appraisal and disposition—what is most important and how to inclusively and ethically represent it. With this in mind in our practice, the robotics community becomes a partner, not just a subject.

The machines and mainframes in our collections challenged us to think differently about our facilities, systems, and processes. We formed new partnerships with our facilities and business offices to develop

workflows and funding models. We formed productive partnerships with marketing and advancement colleagues, leading to shared outreach and communication strategies and several fundraising initiatives. Our colleagues became invested in the overarching vision: they became part of the team.

We recommend collaborating widely with adjacent information specialists.

Libraries, archives, museums, and other cultural heritage areas are working on similar challenges related to digital access and preservation, as well as data management and software curation. Like other STEM fields, robotics generates copious amounts of code, data, and media. We have aimed to develop a thematic collection and, more widely, to design a model and strategy for approaching this amorphous discipline. We are not alone in these efforts and have found partners and allies in historians, digital curators, digitization and metadata experts, scholarly communication librarians, digital humanists, open science specialists, data curators, and other colleagues across our library.

Advocacy & Incentives

Throughout our work on The Robotics Project, we found ourselves explaining—and even justifying—what archivists do and how we can help STEM-oriented researchers. Many individuals we engaged with were curious but consistently unfamiliar with library and archives services. They were receptive but unsure about why they might need or want to work with our team. In short, our investigation revealed that the robotics community did not view us as natural partners.

We recognize that one of the reasons for this uncertainty is a lack of **incentives** for the people who work in robotics. The field is primarily future-facing, and their work is chiefly valued through publications and demonstrations. Many are indifferent to preserving their past work and do not keep consistent records of previous projects.

Although this challenge is not unfamiliar to archivists, especially those who work with the sciences, we arrived at a perplexing insight: building multimodal collections in robotics will likely require a *cultural shift*.

We must help researchers see the value of documenting processes and the benefits of long-term access and preservation. There are clear and urgent opportunities for information professionals and people working in robotics and other scientific disciplines to identify reciprocities and collaborate on solutions for more efficient knowledge management systems and protocols. It would benefit researchers, especially in lab contexts where student and staff turnover is high. Situated within the scientific process, an ***embedded archivist***, bringing together archives and records management practices, could advise researchers and optimize collections for comprehensive and multimodal long-term stewardship.

APPENDIX

SEMI-STRUCTURED INTERVIEW SCRIPT

We will instruct participants not to reveal any private or personally-identifiable information about themselves or others in their responses to our open-ended questions.

Questions

1) What is your role at [insert name of the lab]?
2) Where do you work primarily? Do you work in the lab or at another location?
3) Describe the work of [insert name of the lab].
4) How many years has the lab been active at CMU? If it has been active for many years, how has it changed over time?
5) Approximately how many people work in the lab? Approximately how many current students are affiliated with the lab?
6) Describe one significant robot, project, or product of the lab.
    a) How did the project begin?
    b) How were decisions made throughout the project?
    c) Who was the defined end-user or audience of the project?
    d) How was it made?
    e) What was the workflow of the build process?

f) Were any back-end components (those not viewable/accessible to the robot's intended end-user) custom-made? If so, are these custom components historically significant for future researchers?

g) Describe the internal dependencies of the robot (i.e., knowledge, data, hardware, software).

7) How do you collaborate with colleagues in the lab? How do you collaborate with colleagues in the Robotics Institute?
8) Do you interact with partners outside of the university? If yes, please describe an example.
9) Are there any decommissioned robots, prototypes, or parts in your office, lab, building, or the space where you work? If yes, can you describe them?
10) What challenges do you face in maintaining decommissioned robots?
11) How do you document your work and research (e.g., video recordings, reports)? Where do you store this documentation?
12) What challenges do you face in documenting your work?
13) Are you aware of any records at CMU that document the history of the lab or your area of research? If yes, can you describe them?
14) Are you aware of any records related to your work at high or immediate risk because of factors such as physical location or digital obsolescence? If yes, can you describe them?
15) What do you think is most essential to preserve as part of a robotics archive at CMU?
16) What do you think is the most significant challenge in preserving the history of robotics?
17) Is there anything else you would like to discuss?
18) Is there anything I should have asked you about that I did not?

BIBLIOGRAPHY

Broussard, Meredith. *Artificial Unintelligence : How Computers Misunderstand the World.* Cambridge, MA: MIT Press, 2018.

Bennett, Michelle L., et al. *Collaboration and team science: a field guide*. National Institutes of Health. https://www.cancer.gov/about-nci/organization/crs/research-initiatives/team-science-field-guide/collaboration-team-science-guide.pdf. Accessed on September 25, 2022.

Bruemmer, Bruce, and Sheldon Hochheiser. *The high-technology company: a historical research and archival guide*. Minneapolis, MN: Charles Babbage Institute, Center for the History of Information Processing, University of Minnesota, 1989.

Cunningham, Adrian. "The Archival Management of Personal Records in Electronic Form: Some Suggestions," Archives and Manuscripts 22 (May 1994): 94–105.

Elliott, Clark A. "Understanding Progress as Process : Documentation of the History of Post-War Science and Technology in the United States : Final Report of the Joint Committee on Archives of Science and Technology." Chicago, IL: Society of American Archivists, 1983.

Foti, Petrina. *Collecting and Exhibiting Computer-Based Technology: Expert Curation at the Museums of the Smithsonian Institution*. Abingdon, Oxon: Routledge, 2019.

Franks, Patricia C., ed. *The Handbook of Archival Practice*. Lanham MD: Rowman & Littlefield, 2021.

Haas, Joan K., Helen Willa Samuels, and Barbara Trippel Simmons. *Appraising the Records of Modern Science and Technology : a Guide.* Cambridge: Massachusetts Institute of Technology, 1985.

Hanington, Bruce, and Bella Martin. Universal methods of design expanded and revised: 125 Ways to research complex problems, develop innovative ideas, and design effective solutions. Rockport publishers, 2019.

Harati, Hadi, et al. "Factors affecting the unplanned use behavior of academic libraries users: Towards an axial coding pattern." *Aslib Journal of Information Management* (2019).

Matters in Media Art. http://mattersinmediaart.org/. Accessed February 1, 2023.

National Digital Information Infrastructure And Preservation Program, U.S. *Preserving.exe: Toward a National Strategy for Software Preservation.* Washington, D.C.: National Digital Information Infrastructure and Preservation Program at the Library of Congress, 2013. Web. https://lccn.loc.gov/2013655114.

Noble, Safiya Umoja. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: NYU Press, 2018.

Oral History Association. "Principles and Best Practices." 2018. https://oralhistory.org/principles-and-best-practices-revised-2018/. Accessed on September 27, 2022.

O'Neil, Cathy. *Weapons of Math Destruction : How Big Data Increases Inequality and Threatens Democracy*. New York: Crown, 2016.

Powell, Tim and Sarah Marks. “Science and Technology Archives in the UK: A Paper for the National Archives.” UK: The National Archives, 2016.

Samuels HW. *Varsity Letters : Documenting Modern Colleges and Universities*. Lanham Md. And Chicago Ill: Scarecrow Press and The Society of American Archivists, 1998.

Society of American Archivists. “Crafting Your Elevator Speech.” 2017. https://www2.archivists.org/advocacy/public-awareness/elevator-speech Accessed on September 27, 2022.

Strauss, A., & Corbin, J. *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*. Newbury Park, CA: Sage, 1990.

"Time-Based Media." The Guggenheim Museums and Foundation. Accessed September 28, 2020. https://www.guggenheim.org/conservation/time-based-media.

Vollstedt, Maike, and Sebastian Rezat. "An introduction to grounded theory with a special focus on axial coding and the coding paradigm." *Compendium for early career researchers in mathematics education* 13 (2019): 81-100.

Weideman, Christine. “A New Map for Field Work: Impact of Collections Analysis on the Bentley Historical Library.” *The American Archivist* 54, no. 1 (1991): 54–60.

Yin, Robert K. *Qualitative research from start to finish.* Guilford publications, 2015.

Yun, Audra Eagle and Society of American Archivists. *Archival Accessioning*. Chicago: Society of American Archivists, 2021.