GIVING LIFE TO MEMORY

Manual for Human Rights Civil Society Organizations on Creating and Preserving Digital Archives

GIJTR
Global Initiative for Justice Truth & Reconciliation

International Coalition of
SITES *of* CONSCIENCE

The **International Coalition of Sites of Conscience** (ICSC or the Coalition) is a global network of museums, historic sites and grassroots initiatives dedicated to building a more just and peaceful future through engaging communities in remembering struggles for human rights and addressing their modern repercussions. Founded in 1999, the Coalition now includes more than 300 Sites of Conscience members in 65 countries. The Coalition supports these members through seven regional networks that encourage collaboration and international exchange of knowledge and best practices. The Global Initiative for Justice, Truth and Reconciliation is a flagship program of the Coalition.

www.sitesofconscience.org

Cover: Image shared by CONAVIGUA, GIJTR partner organization in Guatemala.

ABOUT THIS MANUAL

Released in 2023, this publication, *Giving Life to Memory: A Manual for Human Rights Civil Society Organizations on Creating and Preserving Digital Archives*, aims to support civil society actors in creating and utilizing digital archives to support transitional justice mechanisms and advance the broader transitional justice goals of truth-telling, justice and accountability, memorialization, and non-recurrence. The manual combines learnings and recommendations from a two-year project on digital archiving, led by Consortium partners the Humanitarian Law Center (HLC), International Coalition of Sites of Conscience (ICSC), Forensic Anthropology Foundation of Guatemala (FAFG), Public International Law and Policy Group (PILPG) and Documentation Center of Cambodia (DC-Cam), as well as input and expertise generously shared by over 60 civil society organizations who participated in this initiative.

ACKNOWLEDGEMENTS

PRINCIPAL AUTHORS
Elsana Nurkovic and Nenad Golcevski

CONTRIBUTORS
Marc Drouin

CONTRIBUTING EDITORS
Predrag Miletic, HLC
Sarah Case, ICSC
Dario Colmenares, ICSC
Bethany Houghton, PILPG
Marc Drouin, FAFG

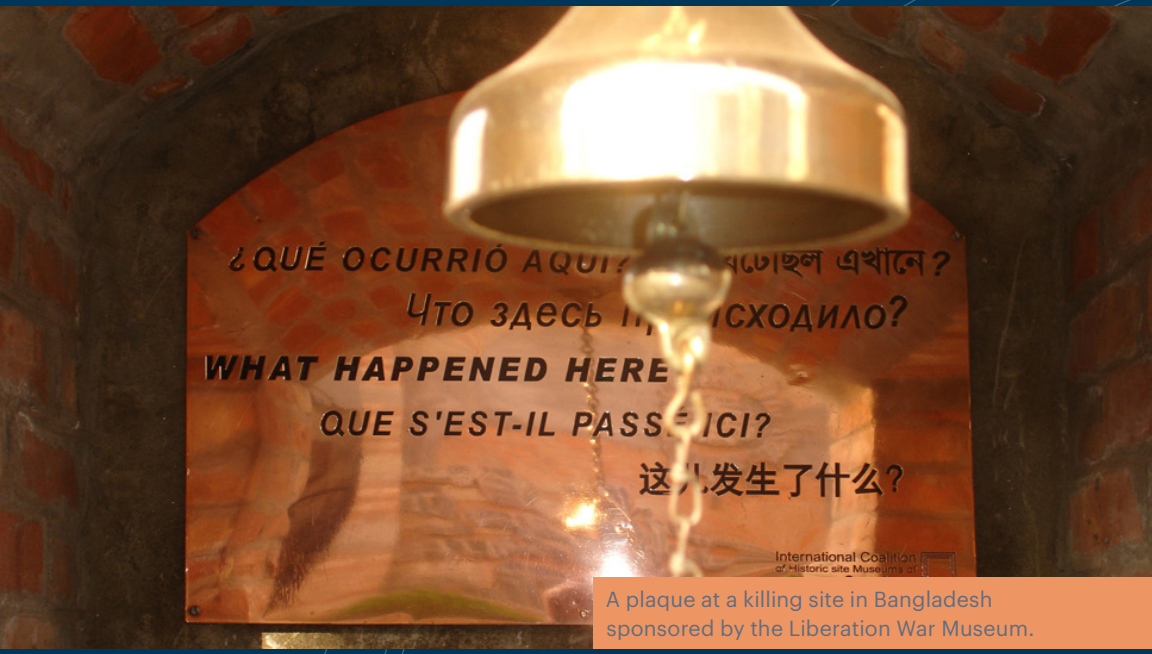
In addition, the Consortium partners would like to express their deep appreciation to the civil society organizations who contributed their time and insights to this manual by sharing challenges, lessons learned and best practices from their experiences in digital archiving. A list of these organizations can be found in Addendum V. The Consortium partners would like to express their particular thanks to the four civil society organizations who helped to pilot and refine an earlier draft of this manual: l'Association des victimes, parent et amis du 28 septembre 2009 (Guinea), La Coordinadora Nacional de Viudas de Guatemala (Guatemala), La Asociación de Familiares de Detenidos-Desaparecidos de Guatemala (Guatemela), and Comision Colombiana de Juristas (Colombia).

Learn more at www.gijtr.org

ABOUT THE GLOBAL INITIATIVE FOR JUSTICE, TRUTH AND RECONCILIATION CONSORTIUM

Around the world, there is an increasing call for justice, truth and reconciliation in countries where legacies of grave human rights violations cast a shadow on transitions. To meet this need, the International Coalition of Sites of Conscience (ICSC) launched the Global Initiative for Justice, Truth and Reconciliation (GIJTR) in August 2014. The goal of GIJTR is to address new challenges in countries in conflict or transition that are struggling with their legacies of past or ongoing grave human rights violations.

The GIJTR Consortium ("the Consortium") comprises the following nine partner organizations:



A plaque at a killing site in Bangladesh sponsored by the Liberation War Museum.

- International Coalition of Sites of Conscience, in the United States (lead partner);
- American Bar Association Rule of Law Initiative (ABA ROLI), in the United States;
- Asia Justice and Rights (AJAR), in Indonesia;
- Centre for the Study of Violence and Reconciliation (CSVR), in South Africa;
- Documentation Center of Cambodia (DC-Cam), in Cambodia;
- Due Process of Law Foundation (DPLF), in the United States;
- Forensic Anthropology Foundation of Guatemala (Fundación de Antropología Forense de Guatemala – FAFG), in Guatemala;
- Humanitarian Law Center (HLC), in Serbia; and
- Public International Law & Policy Group (PILPG), in the United States

In addition to leveraging the different areas of expertise of the Consortium partners, the ICSC draws on the knowledge and longstanding community connections of its 275-plus members in 65 countries to strengthen and broaden the Consortium's work.

The Consortium partners, along with the ICSC's network members, develop and implement a range of rapid response and high-impact programs, using both restorative and retributive approaches to criminal justice and accountability for grave human rights violations. The Consortium takes an interdisciplinary approach to justice, truth and accountability. On the whole, the Consortium partners possess expertise in the following areas:

- Truth telling, memorialization and other forms of historical memory and reconciliation;
- Documenting human rights violations for transitional justice purposes;
- Forensic analysis and other efforts related to missing or disappeared persons;
- Advocating for victims, including for their right to access justice, psychosocial support and trauma mitigation activities;
- Providing technical assistance to and building the capacity of civil society activists and organizations to promote and engage with transitional justice processes;
- Reparative justice initiatives; and
- Ensuring and integrating gender justice into these and all other transitional justice processes.

Given the diversity of experiences, knowledge and skills within the Consortium and the ICSC's network members, the Consortium's programming offers post-conflict countries and countries emerging from repressive regimes a unique opportunity to address transitional justice needs in a timely manner while simultaneously promoting local participation and building the capacity of community partners.

A 2018 GIJTR exhibition of body-maps in Conakry, Guinea.

CONTENTS

PART 1

INTRODUCTION

"It is all there. All the documents have been in that big suitcase in my basement for the past 30 years – the letters, petitions, photographs, press clippings… They could help explain what actually happened back then and how" – a friend, human rights activists, told us recently.

And yet, we all agreed, as long as those documents stay in that suitcase, in that basement, they are virtually nowhere. They are not helping explain or teaching our, or other societies, any of the many things they could. Moreover, they are in grave danger of being lost, as they degrade over time.

Many of you reading this will have encountered such suitcases – or perhaps card boxes or overcrowded shelves – filled with invaluable historical material about events, crimes, victims, legal processes, or the wider context of human rights abuses committed during an armed conflict, or a violent regime rule. Or perhaps you might have stumbled upon a box of hard drives, or stacks of DVDs, with seemingly innumerable digital files with witness statements or audio-visual recordings testifying to police brutality and citizens' protests.

That kind of material is of immense social and historical value, and when faced with it – say a statement of a survivor of a mass execution or a letter of a husband to his wife from a concentration camp or a scarf made by a mother for her disappeared son – no reasonable citizen, let alone a human rights activist, could possibly allow it to be lost and forgotten. Our core values, our very instinct, all our training, and our experience scream at us that such invaluable evidence, such pieces of history, must be preserved – as well as made public – for the current and future generations of citizens dealing with the legacy of violence from their country's past.

Many of you understand very well the value of such material, as you might have used it to help victims realize their right to reparation; to identify perpetrators of crimes and bring them to justice; or to support truth-telling, memorialization, or education about the violent past. What comes along with such experience is the understanding that it is not sufficient to merely possess certain documents or material, no matter how valuable their content.

Image shared by AVIPA, GIJTR partner organization in Guinea.

For such material to be utilized for any of the noble causes of human rights defending or for transitional justice, we need to know exactly what it contains, structure and describe it, identify each item within it, define its relationship with other items or groups, and store it in a way that will allow us, and others, to find it again later and be able to safely retrieve and use it.

In other words, we need to create an archive.

But this archive should also be easily searchable, with simple access to each item and its description. The public should have access to nonsensitive material and should be able to copy, analyze, and reuse it for legal, educational, artistic, or other purposes. More immediately, the archive's contents should be safely preserved, possibly for the long term, while backup copies should be created and the entire content should be stored safely, with the possibility to move it elsewhere if needed.

In other words, we need to create a digital archive.

Put plainly, unless the gathered records of crimes and human rights abuses (physical and digital alike) are preserved in a planned, organized, and sustainable digital form and system, their immediate usability will be very limited – and in the long term they are unlikely to be preserved, let alone leveraged to fully serve their purpose in the future. This is why the need for digital archiving goes to the very core of the hard, painstaking work of numerous organizations worldwide that record and safeguard material on human rights violations.

And this is why this manual was created: to guide the Civil Society Organizations (CSOs) through this process and assist them in creating and developing their own digital archives.

PART 2

MANUAL OVERVIEW

2.1 Basic Concepts

The purpose of this manual is to be an instructional and practical guide to digital archiving for CSOs.

Digital archiving is the process of creating, managing, and developing digital archives. For the purposes of this manual, digital archives are defined as archives that contain material in a digital form – including both born-digital items and digitized versions of originally physical material – stored on digital media and managed through digital tools.

It is important to underline, however, that digital archives are still archives that contain documents or other material, albeit in a digital form, that needs to be organized, described, arranged, stored, safeguarded, and made accessible to users. Therefore, in many ways, digital archives are just like any physical archive in that they require application of the same sets of basic archival rules, techniques, and actions. In this sense, there is a significant overlap between the Guiding Principles, methods, and work needed to create both a physical and a digital archive.

Image shared by CONAVIGUA, GIJTR partner organization in Guatemala.

Nevertheless, archival work with digital material, tools, and resources also brings a new layer of requirements, considerations, and challenges. From planning the archive and organizing the material to decisions on its description and arrangement, use of specific software, managing storage, and defining modes of access, digital archiving adds a different spin to the archival process. It requires a somewhat different approach and an additional set of archival techniques. It is precisely these specific aspects of digital archiving work that this manual addresses, outlines, and provides guidance for.

The following are reasons why the focus of this manual is specifically on digital archiving.

2.2 Why a Manual on Digital Archiving?

Born-digital and digitized physical materials have become widely present in archives created and developed by CSOs dealing with human rights abuses. Moreover, physical archival material can be preserved and secured only through digitization and by making digital copies a part of the digital archives. Another growing need for digital archives arises from the often-excessive amount of material that needs to be archived and preserved long-term, thus the only feasible way of storing and preserving it is through digitization and the creation of a digital archive.

Digital archives can also provide immensely improved internal and external access to content and the possibility to search, identify, manage, and review or copy any item in the collection. This, in turn, opens completely new horizons – in terms of the potential for digital archival material to be used for outreach or educational or memorialization purposes.

If we are to successfully preserve the memory of the past by creating archives today, we must always bear in mind their future users. We should be mindful that the archival work and solutions we apply need to be communicable and transferable to those that will use it in the future so that our archives can be sustained long-term. To ensure this “future-proof” quality of our archives, it is necessary to use tested and proven digital tools, formats, and resources for managing and archiving both born-digital and digitized archival material.

Image shared by Conavigua, GIJTR partner organization in Guatemala.

For all these reasons, among CSOs working with material on human rights violations there is a concrete and growing need for information and resources on digital archiving – including a manual such as this one. On the one hand, digital archiving of such material is quickly becoming a necessity for these organizations. On the other hand, CSOs often lack the necessary knowledge and skills for creation and development of digital archives, and would therefore benefit from a resource that provides structured guidance through the process and an overview of challenges, tips, and best practices. While there are already very good Manuals available dealing with different aspects of physical archiving, there is no systematized, yet practice-based and practice-oriented guide for CSOs, addressing the entire process of digital archiving.

This gap has been identified by the Global Initiative for Justice, Truth and Reconciliation (GIJTR) – a consortium of nine global organizations that work together to assist communities in, or are emerging from, conflict in creating just and peaceful futures. To address the growing need of CSOs for information resources, capacity building, and networking in the field of digital archiving, in 2021, five organizations – members of the GIJTR consortium – initiated the project "Supporting CSOs in Digital Archiving."

This manual is a result of this project. It was conceptualized, devised, and developed on the basis of research, exchanges, and knowledge gathered throughout this two-year project. It is grounded in, and built on, the experiences and insights from more than 40 CSOs from 24 different countries that provided their input and participated in this project.

2.3 Framework for the Manual: "Supporting CSOs in Digital Archiving" Project

The "Supporting CSOs in Digital Archiving" project was designed to aid civil society actors in creating and utilizing digital archives to support transitional justice mechanisms and advance the broader transitional justice goals of truth-telling, justice and accountability, memorialization, and non-recurrence.

The project was implemented in two stages, over two years, by five organizations/members of the GIJTR consortium. The project leader was the Humanitarian Law Centre (HLC) from Serbia and the consortium included the International Coalition of Sites of Conscience (ICSC) from USA, Public International Law & Policy Group (PILPG) from USA, Documentation Center of Cambodia (DC-Cam), and Forensic Anthropology Foundation of Guatemala (FAFG).

The project's goals were to increase understanding and knowledge of the digital archiving-related challenges, needs, and practices of CSOs dealing with mass human rights violations; to facilitate networking and exchange of experiences and lessons learned between them; to increase their capacity for digital archiving; and to develop practical resources on digital archiving that respond to the CSOs' varied needs.

To these aims, GIJTR partners designed and disseminated two questionnaires: one aimed at assessing the digital archiving-related needs and practices of CSOs that are only starting to develop their digital archives, and the other to gather information on best practices as well as challenges encountered by CSOs that are already advanced in the process of digital archiving. Responses to these two questionnaires, either in written form or through interviews, have been collected from more than 40 relevant CSOs spanning three continents – from Japan to Croatia.[1]

Image shared by AVIPA, GIJTR partner organization in Guinea.

The first project phase also included a three-day virtual dialogue on digital archiving needs and best practices for CSOs, with participation of GIJTR consortium members as well as many of the organizations that provided their input through questionnaires and interviews. During this phase, the partners also conducted a review of the existing literature on digital archiving for CSOs and gathered and made readily available the identified relevant resources.

In the second stage of the project, relying on the analysis of, and findings from, the questionnaires and interviews with CSOs as well as the insights gathered from the exchanges held as part of the project's virtual dialogue, GIJTR consortium members developed a framework for this manual. Hence, the experiences, practices, challenges, and needs of CSOs with hands-on experience in digital archiving make up the foundation of this manual, are its main source of material, and guided the selection of topics and issues it addresses.

Moreover, a draft version of this manual was used by four selected CSOs to guide implementation of their respective pilot projects in digital archiving: "The Association of Relatives and Friends of the Events of September 28, 2009" (AVIPA, Guinea); "The National Coordination of Widows of Guatemala" (CONAVIGUA); "The Association of relatives of the detained and disappeared of Guatemala" (FAMDEGUA); and "Colombian Commission of Jurists" (CCJ).

Following a six-month period of pilot project implementation and the use of a draft of this manual as a guide, the feedback provided by these organizations has been incorporated and used to improve the final version of the manual to make it as useful as possible to CSOs in their daily digital archiving work.

2.4 Purpose and Structure of the Manual

As the purpose of this manual is to be an informative and practical guide to digital archiving and creating digital archives, it relies on examples and best practices collected from GIJTR Consortium members and partners. It is structured in a way to make it useful to a wide scope of human rights organizations, both those that are only beginning this process as well as CSOs working to maintain and further develop their digital archives.

Image shared by CCJ, GIJTR partner organization in Colombia.

The manual combines current expert knowledge and insights from the practice of digital archiving with the perspective of human rights organizations around the world and their specific needs, challenges, and best practices in this field. It draws on the direct experiences of CSOs that took part in the project and shared their views through detailed questionnaires, exchanges, and discussions.

Starting from the decision to create a Digital Archiving System through its establishment and development to its continuous change, adaptation, and maintenance, the manual acts as a guide throughout the process. Various organizations reading and using it will be at different points of their digital archive life cycles. Therefore, while the manual is written to assist through the entire life cycle, it is also designed to allow for using separate chapters as standalone sources for informing specific phases of digital archiving work.

PART 3

VALUE OF DIGITAL ARCHIVING FOR CIVIL SOCIETY ORGANIZATIONS

The trouble with civil society-collected archival material on mass human rights violations is it was never intended to be an archive.

CSOs collect documents, photographs, audio-visual recordings, press clippings, artifacts, and other materials not in order to archive and preserve them but rather to achieve concrete impact on a specific legal, political, or social process and event. An organization might be collecting signatures of citizens opposing a war to petition the government to stop the conflict, or a CSO might be gathering testimonies of war crimes survivors and witnesses to provide as evidence in a legal case against certain perpetrators, or for providing reparations to victims. In any case, the focus of CSO activists working in a time of conflict or a violent regime rule will, very reasonably, not be on identifying and listing each item and carefully describing, arranging, and structuring the gathered material. Rather, they will be dealing with overwhelming events at hand, trying their best to just keep working and collecting as much material as possible.

CSOs collect data on a wide scope of human rights violations, including killings, disappearances, imprisonment, torture, gender-based crimes, and many others. These records are often complemented by documents relating to relevant publications; studies; documentation on political and public events; public perception of various issues; and data on social and economic impacts of authoritarian regimes and armed conflicts. Further, these materials can also include records of activities of different political, social, or armed groups, as well as CSOs.

Other types of data include documents relating to peace-building and transitional-justice processes, such as judicial proceedings, truth-telling sessions, commemorative practices, and reparation programs. These records can also include statements from survivors, witnesses or victims' families, various judicial documents, analyses, media reports, photographs and audio-visual material, as well as items and artifacts such as campaign materials, diaries, drawings, letters, and the like.

As a result of the context and manner in which they have been gathered by CSOs, such collections can often be found as groups of nonstructured material – perhaps as stacks of folders with paper documents or groups of unmarked

audio-visual files. The actual size of the collection might be unclear, along with the categories and types of material it contains. Sometimes, such material can be in danger of physical degradation or other types of harm.

Yet, as the amount of collected material continues to grow, along with the need to search and access it, it becomes increasingly clear that the collection has, in addition to its operational purpose, gained historical, legal, and social value. This is often the case once a conflict or a violent regime rule has ended and that society enters transitional justice processes. The CSOs' archival collections can be used for achieving accountability, supporting reparation programs, developing memorialization initiatives, and devising educational resources.

For CSOs that have gathered or obtained such invaluable collections of materials that are – in archival terms – undefined, unorganized, and therefore unusable, transforming these collections into proper, usable, and sustainable archival collections becomes a necessity. When these archives contain physical material that needs to be digitized, born-digital, or both, achieving this goal will require an organization to embark on the process of digital archiving.

Creating a digital archive can ensure long-term preservation of the collected material, especially if it might be at risk. By digitizing physical collections, we are creating copies that can be sustainably preserved independently of the original item and its condition. Moreover, for both digitized physical items and born-digital files, creation of a digital archive allows us to ensure that there are sufficient backup copies of the archival material – and that they are safely stored and easily located to ensure their security and resilience. In this way, digital archiving allows us to retrieve the archive's content even in cases when the material in the main collection has been lost, corrupted, or destroyed.

Beyond mere preservation, creation of a digital archive allows for substantial enhancement in terms of managing and operating the content, as well as identifying, searching, locating, and accessing its individual items. When it comes to content management, digital archiving is a true "game changer" in terms of the number and scope of improvements it allows for. Starting from the obvious, rather than going through shelves or boxes, the archivist can click through a database and move an item from one collection to another, or change its name or description. Content can be listed, reviewed, retrieved, copied, or checked for errors – or virtually any other archival action can be performed more easily and efficiently. An important additional benefit of digital archives is they allow for each

Image shared by FAMDEGUA, GIJTR partner organization in Guatemala.

action taken on any archival item to be recorded. Hence, in addition to an item itself, the full record of actions taken on it is also preserved in the archive. This is important in determining the chain of custody over an item and is therefore of particular relevance for ensuring credibility of archival content.

An equally significant advantage of digital archiving – in particular for CSOs working with archival material on human rights violations – is the immense improvements it brings to the potential for external access, outreach, and use in education, as well its leverage in various processes of transitional justice. Digital archives containing nonsensitive public material can be made easily accessible online to anyone, anywhere in the world – not just mere access to the material, but also the quality of that access is substantially advanced, as digital archives enable meaningful search of content using various criteria; review and analysis of selected groups of items; simultaneous access and use of material by multiple users; and copying of individual files by external users.

Such democratized and enhanced quality and experience of access and use of digital archival material enables its more extensive and varied use – by not only external users but also the CSOs that create and own those archives. This is because digitally archived material can efficiently, and in a variety of formats, be made available, presented to the public, or used for content production, education, and many other purposes. The CSOs can – and do – raise awareness of their archives, and thereby extend their visibility, reach, leverage, and impact.

The digital archives created during a time of conflict or violent regime rule can be used for efforts to achieve peace, support victims and potential targets of violence, and enhance democratic processes. More often, the digital archives are developed and leveraged in post-conflict and post-authoritarian periods as important tools for transitional justice processes, as they are used to inform and teach about the violent past, fight impunity, and support truth-telling, memorialization, and reparation programs. This often includes development of educational programs and resources, establishment of museums and documentation centers, as well as support for projects aimed at improving the rights of victims or advancing the processes of reconciliation. Digital archives are also being used for academic research, media reporting, and artistic projects dealing with the violent past.

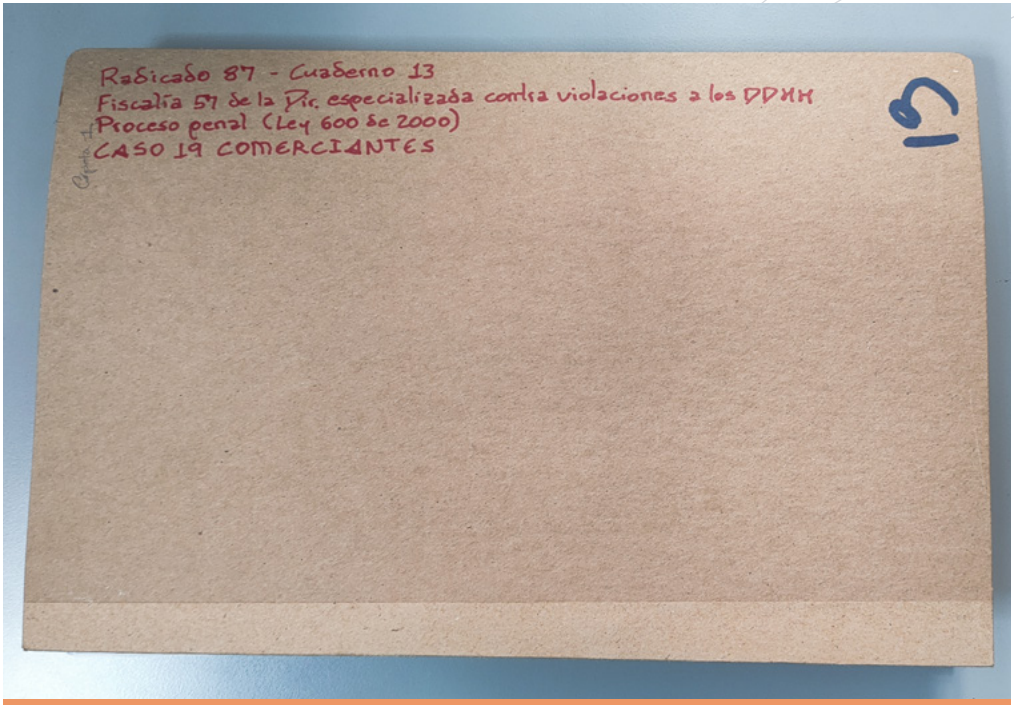


Image shared by CCJ, GIJTR partner organization in Colombia.

These various uses of digital archives can and are being implemented by not only CSOs themselves but also by other, external users: victims' groups, the media, legal professionals, academics, artists, etc. Digital archives streamline access and allow for a wider visibility into their content, thus creating a multiplier effect that significantly extends their reach, raises the extent of their use and the number and scope of users, and ultimately strengthens the archives' overall impact.

PART 4

MAIN CHALLENGES FOR CSOs CREATING DIGITAL ARCHIVES

The trouble with CSOs building digital archives is they never expected to become holders of archives.

CSOs usually have neither resources and capacities nor the expertise needed to create digital archives. Yet, the value of the material they have gathered and the need to preserve and leverage it is so high that organizations will often have no choice but to take on this unexpected role of digital archivists.

For many CSOs, this can be a challenging and trying experience, as digital archiving is a complex and demanding function that requires significant human, financial, technological, and time resources – something CSOs typically do not have in abundance. Members of CSOs in such a position often speak of utter frustration and the feeling of being overwhelmed with the number and complexity of requirements involved with building a proper digital archive without having the appropriate means to address them.

The key challenge in this phase, as well as in the entire process of digital archiving, is to persist in spite of obstacles, not to give up, and patiently find a way to deal with the many difficulties that will inevitably come. The holy grail of preserving and leveraging invaluable, painstakingly gathered material that would otherwise be lost must certainly outweigh the high level of difficulty of the task itself.

There are indeed many inspiring stories of CSOs that faced seemingly insurmountable challenges yet managed to overcome or find their way around them and eventually build successful digital archives.

Initial challenges an organization faces will likely be internal, related to the organization itself and its functioning. This includes the frequent lack of capacities needed for developing a digital archive – including expertise in digital archiving, human resources, financial means, necessary equipment, infrastructure, and space. But there are also further difficulties that often occur later in the process of developing a digital archive.

In relation to finances, the difficulties often involve lack of donor interest to support the long and resource-demanding process of building a digital archive. Moreover, building digital archives can involve unexpected costs and slower-than-expected progress, which makes it even more difficult to raise the necessary funds.

Image shared by FAMDEGUA, GIJTR partner organization in Guatemala.

Handling human resource demands is equally challenging for an organization as building a digital archive. First, there is a need to significantly increase the number of staff, as digital archiving – especially in its early phases – requires a substantial workforce. This means an organization needs to quickly adapt to its new size, which can be difficult if the increase is substantial. Further, this often brings along a high fluctuation of archival staff with specific, sometimes unique knowledge of the archives, which creates two types of difficulties: it can lead to loss of institutional knowledge and it creates additional burden in terms of the need for constant training of new staff members. A separate, important challenge for most organizations is obtaining and keeping staff with the necessary expertise – in the archival or technical domain – which is needed throughout the process of building and developing a digital archive.

In cases where building a digital archive involves digitization of collections of physical materials, additional challenges are involved. Beyond obtaining the necessary equipment, software, and other technical solutions, as well as the necessary human resources, an organization also needs to address the issue of long-term preservation and storage of the original, physical materials, which can require further resources, time, and expertise.

The quick pace of technological development and the stream of ever-new solutions in the field of digital archiving presents another challenge for CSOs. It can become difficult for an organization to simply follow the latest trends and solutions, let alone be able to obtain them, train its staff to use them, and implement them. However, sometimes the changes and updates are necessary as formats and technologies become obsolete. These can be especially trying challenges for CSOs developing their digital archives, as migration of archival data or transition to new software or hardware requires specific and extensive financial resources, time, workforce, and expertise. Many organizations also report that lack of adequate internet infrastructure in their country – primarily a stable internet broadband connection – creates significant challenges for the development of their digital archives.

Another frequent concern for CSOs in relation to digital archiving is the security and safety of the digital archival material. This relates to physical and cyber-threats to data security, and also involves the challenges of ensuring that sensitive and confidential data contained in the archive is not leaked to the public or to an unauthorized person or organization.

Image shared by CONAVIGUA, GIJTR partner organization in Guatemala.

Finally, CSOs building digital archives often report that an important challenge in the process is an organization's lack of adequate expertise. Hence, the organizations frequently need assistance in the form of external expertise for guidance or mentoring throughout the process. Simultaneously, CSOs have a need to build their own capacities and grow their internal experts and institutional knowledge through practice and regular, targeted training of their members.

On top of this extensive set of internal, organization-related difficulties being insufficient, CSOs developing digital archives also regularly face several external challenges. These can include political pressure or security threats to the organization or its archives from people and organizations that do not want the archival material to be made public and accessible or used in legal processes. This is frequently the case for CSOs that build their digital archives in a time of conflict or violent regime rule, or in post-conflict and post-dictatorship periods in countries where elements of previous regimes – individuals or organizations – have retained substantial power and media and social control.

Often, the work of organizations documenting human rights abuses is condemned publicly by politicians and media that support them, with their digital archives proclaimed inaccurate or false. This can lead to a CSO's negative public perception and alienation from its own community, which compromises staff well-being and security and can create practical difficulties for an archive's development and daily operations.

After reading through this extensive-yet-not-even-comprehensive thread of difficulties and obstacles involved with digital archiving for CSOs, you would be forgiven for asking why anyone would ever wish to bring upon their organization all these seemingly insurmountable challenges, and how an organization could successfully navigate them.

The answer to the first part is simple, but worth repeating: The difficulties along the process of creating a digital archive are many and require patience, perseverance, and readiness for the organization to take some blows, while progress should be measured in small steps. Yet, given that the goal of preserving and, indeed, giving life to memory is so important and valuable, that trade-off is clear: This seriously challenging process of digital archiving is well worth taking for any CSO with a valuable archive to preserve.

The answer to the second part of the question is contained in the chapters that follow, which address the various challenges encountered by CSOs in the process of digital archiving. The core purpose of this manual is, as mentioned, to assist organizations in addressing these challenges – which are discussed throughout this manual in the context of the specific stages of digital archiving to which they are related – in their daily work with digital archives.

PART 5

DIGITAL ARCHIVE LIFE CYCLE

The trouble with digital archiving is it is not really archiving – or at least not only archiving.

Rather, digital archiving is a never-ending process of transformation of the digital content one is trying to save from oblivion, and of the system in which it is preserved.

In that sense, digital archiving is a lot like that famous line from the song *Hotel California*: "You can check in any time you like, but you can never leave."

Digital archiving is not a process that ends at a certain point. Any content we enter into a digital archive – any solution we apply to its storage, preservation, security, or access – is bound to eventually be transformed, the data migrated, and the technologies replaced.

To borrow a metaphor: Let us compare digital archiving to archiving of an object, say an antique, 5,000-year-old clay tablet. To preserve that physical object, we can leave it to sit in its storage space and only need to ensure that the optimal conditions in which it is stored are *not changed*. The opposite is true with digital archiving objects: To preserve them, we have to *continuously change* digital objects and their environment. We have to migrate data and transform the archival system to avoid format, storage, software, or other technologies becoming obsolete.

Digital archiving therefore does not have an end point. Rather, it is a cyclical process in which stages follow one after the other continuously, without a final destination. Creation of a digital archive can hence be seen as only the beginning of the process – the cycle's first iteration – which will then be repeated for as long as we wish to preserve the archive's digital content.

To reflect this fundamentally important characteristic of digital archiving, its circular and continual character, the manual applies the "**Digital Archiving Life Cycle Model**."

The Digital Archiving Life Cycle Model also usefully presents and makes salient several other key characteristics of the digital archives. It draws attention to the need for taking actions and actively managing a digital archive throughout its life cycle. At the same time, the model clearly presents the wide scope of

responsibilities involved in the digital archiving process. Finally, the Life Cycle Model makes clear and salient the important fact that decisions and actions in each phase affect what can be done and how in each following stage and any new iteration of the process.

Having awareness of these dynamic relationships between all the phases in digital archiving is needed to make informed decisions in each phase so that they do not limit the possibilities for actions and solutions in the subsequent stages.

It should be said that there is no single universal model to describe the digital archiving process. The models applied vary depending on an archive's content, purpose, and users, as well as the archiving organization's policies and practices. The Life Cycle Model of digital archiving used in this manual was developed to tailor to the specific needs and challenges of CSOs. It reflects some of the elements of the OAIS Reference Model and partly the structure of the DCC Curation Life Cycle Model. The OAIS Model is the most widely used model for digital archiving, while the DCC Life Cycle Model includes many of the considerations that also affect CSOs engaged in digital archiving.

Image shared by CCJ, GIJTR partner organization in Colombia.

This manual applies a simplified Life Cycle Model that focuses on key aspects of the process for CSOs. Presented visually, the model shows the main stages of digital archiving following each other in the shape of a circle, just as numbers do on a clock, with the end point marking the beginning of a new circle – a new, slightly different iteration of the process. See Figure 1.

Figure 1. Digital Archiving Life Cycle Model

0:00

Once a strong need for creation of a digital archive has been identified and a firm organizational decision has been made to develop it, the process begins with the **Planning and Organizing stage**.

As a first step, we need to develop a General Plan, which will define the Guiding Principles of the archive as well as address key organizational, technological, and resource-related issues that will be encountered throughout the digital archive's life cycle. The Guiding Principles are based on responses the organization gives to a set of core questions, such as, What needs to be preserved? Why? Who will use it? And how?

The General Plan needs to be complemented with the *creation of an Identification Inventory, selection, organization, and description of the material we want to preserve*. This is because any further decision or action in the process will rely on information about the format, amount, scope, size, topic, or other characteristics of the collected material for preservation, as well as the ability to identify, manage, and locate groups or individual items.

To round off this stage, we will need to *plan, design, and select our future Digital Archiving System – a digital repository and content management system* that will host our archival content. A Digital Archiving System consists of hardware and software elements that will need to be carefully selected, given that their characteristics will affect other important aspects of our digital archive.

The Planning and Organizing stage is the foundation for the creation of any archive, including a digital one. It shapes all the other stages and defines the decisions and actions to be taken in them. Different elements of the Planning and Organizing stage will need to be revisited, consulted, and reviewed at various points later in the process. Finally, at the close of a digital archive's life cycle, the process will return to this initial phase, this time to plan and organize for a digital archive's development and transformation through the next iteration of its life cycle.

3:00

The second stage includes a group of **Digitization, Description, Preparation, and Preservation** actions that lead to the process of inputting *our digital material into a Digital Archiving System***.** These actions are separate but go hand in hand, as they are interrelated and need to be well-coordinated. Digitization of any physical material needs to be done in sync with the decisions regarding how these objects will be described (i.e., which information, or metadata, about them needs to be captured in the digitization process) – much like born-digital material, whose metadata needs to be selected as well.

This is followed by a number of actions aimed at proper preservation of the archive's content by maintaining its integrity and credibility (i.e., ensuring that the objects are not compromised and that any changes made to them are recorded).

The material, both digitized and born-digital, is then ingested into the Digital Archiving System and onto the storage media. In this process, the content and its descriptions – its metadata – are captured and stored in the Digital Archiving System. Additional checkups are then performed and backup copies are created and stored separately.

6:00

Providing Access and Data Security is the main task in the third stage of digital archiving. These two separate functions are interrelated and need to be kept in balance to provide for the optimal effect – the widest possible access to be provided – while maintaining data safety and protection. This includes both protection of any private, sensitive, or copyrighted data and measures to provide for safety of data and storage systems for protection against physical harm and cyber-threats.

Providing wider access, for example by making a digital archive accessible through open databases or online platforms, will pose an additional set of data security issues compared with providing access

to a closed group of users. Similarly, different items in the digital archive may require varying levels of protection and controlled access. Therefore, appropriate levels of access need to be defined for different groups of users in relation to different parts of the archive.

9:00

Maintenance Through Preservation and Migrations is the action that dominate the fourth stage of digital archiving. Once the digital archive has been designed, set up, and populated – and its data preserved, secured, and made accessible – all these functions need to be maintained and monitored and the content and the system need to be managed and eventually migrated and transformed.

Regular maintenance checks need to be performed on the data (to ensure its continued integrity and credibility, as well as format usability), the system (to provide for continued security and open access), and hardware and software technologies (to ensure their proper functioning and act timely when they need to be migrated or transformed to prevent them from becoming obsolete).

0:00

At this point, another iteration of the digital archiving process begins anew.

PART 6

0:00

PLANNING AND ORGANIZING

The importance of thorough and careful planning and organizing of a digital archive at the very beginning of the process cannot be overestimated. A well-devised plan for the archive will provide the grounds and guidance for decisions and actions throughout the digital archiving process. On the flip side, a poorly considered decision or an omission in this phase will create additional difficulties in further phases and actions in a digital archive's life cycle.

The key activities at this stage include devising a General Plan for the digital archive; creating an inventory and selection of the material for preservation; organizing and describing the material to devise a structure for the future archive; and planning the Digital Archiving System and selecting its main hardware and software components.

6.1 General Plan

Creating the General Plan is a crucial first step in the process of developing a digital archive. It lays out the reasons for and the method of the archive's development by providing it with Guiding Principles as well as key decisions regarding the content, access, and major organizational, technological, and resource-related issues. Such widely scoped, detailed, and advanced planning will help the organization navigate a wide array of challenges that will need to be met in the later stages of the process of digital archive creation.

It is important to note that the General Plan should record not only the conclusions and decisions but also the reasoning and grounds on which they are made, as doing so assists their later review and potential revision, especially when context or circumstances change.

There is no universal template for a General Plan for digital archiving, and this document's usefulness will differ somewhat depending on the content and context of the collection, as well as the organization itself. However, there is a

Image shared by CONAVIGUA, GIJTR partner organization in Guatemala.

set of questions that can be a useful guide for development of a General Plan. These questions concern the content and purpose of the future archive, as well as organizational, technological, and resource-related issues. Providing detailed, well-informed, and considered responses to this set of questions will give you a solid basis for devising a General Plan.

Figure 2 provides an example of a list of questions that should be answered in developing a General Plan for a digital archive. Please note that this is just an example of a list, not a template, and as such it can be amended and tailored to the needs of a particular archive and organization.

General	Content and Format	Resources
What is the purpose of the archive you are creating? Why are you setting it up? What do you hope to achieve with it? Who are intended users?	What restrictions exist on access to items? Are there legal or moral considerations on data privacy? What collections make up the archive? Will the archive host items in any format?	What time and resources will be needed to develop the digital archive? Can you build capacities and develop the human resources? Can you raise funds or find other means to obtain software and hardware?

Figure 2. Example of a list of questions to be answered in a developing General Plan for a digital archive

The responses to these questions can be divided into different segments of the General Plan. These will serve as the Guiding Principles for the development of the digital archive.

Guiding Principles

The Guiding Principles summarize the reasoning behind the development of a digital archive. They state why an archive is needed, who will be using it and how, and what the expected benefits of its creation and development will be. The Guiding Principles also address several other key issues, including the required resources and technologies, legal and security-related responsibilities, and organizational matters.

The Guiding Principles should serve as a reference point, a measuring stick for any future major decision or action to be taken in the process. For example, if one of the benefits of the digital archive is maintaining credibility of data and recording the chain of custody over a digital object, then we can rule out any software or system solution that does not perform well on that function. Similarly, we will not implement any data security solutions that would obstruct access for a key group of users.

Although foundational for a digital archive, the Guiding Principles are not "set in stone" and can and should be reviewed and amended when necessary. Over time, with changes in the archive's external community, technological development, and the iterative transformation of the archive itself, the organization might decide to alter the archive's Guiding Principles to better suit them to the changed environment.

A fictional example of a General Plan is provided as an addendum at the end of the manual. In that example, we include a set of Guiding Principles that should be included in any digital archive's planning and development. It also provides brief descriptions of main considerations and issues to be addressed by each Guiding Principle, and how these can be formulated. This example should not be considered a definitive list of Guiding Principles, nor used as a template.

6.2 Identification, Selection, and Prioritization

Simultaneously with the development of the General Plan, we need to identify, evaluate, organize, and describe the material we wish to preserve. This will allow us to map the material and gather and arrange key information on its

Image shared by FAMDEGUA, GIJTR partner organization in Guatemala.

characteristics, which creates the basis for further archival processing. It is also a necessary step to enable us to make any planning and decision-making on how the material can be archived and preserved and how the Digital Archiving System can be built.

Handling Unstructured Physical Archival Material

At this point, many CSOs following this manual in developing their digital archive will find themselves faced with the challenge of handling numerous, unorganized batches of their physical material – be it boxes full of mixed-up files, shelves with random folders and documents, or boxes full of unmarked VHS tapes.

The difficulty these organizations encounter is how to properly deal with such physical material and turn it into organized, labeled, and safely preserved physical archival content, which could only then be digitally archived.

We know this to be a common situation among CSOs – potential users of this manual – based on substantial and consistent feedback we have received to

that effect. Thanks to the unique design of GIJTR's project "Supporting CSOs in Digital Archiving," this manual has a benefit of having been piloted by four CSOs and then reviewed by a wider group of relevant CSOs that provided their comments and recommendations.

Much of this feedback clearly pointed to a need for detailed, hands-on instructions on how to properly approach, handle, organize, and ensure long-time preservation of unstructured physical material the CSOs wish to archive before the digital archiving process can begin.

Additionally, the need for this type of practical guide came from another line of CSO-provided feedback, stressing that the manual should lay out in more detail and in hands-on form the necessary archival procedures and concrete tasks CSOs need to take in organizing, describing, and preserving the physical material as a precondition to its digital archiving.

To a great benefit of this manual and its future users, precisely such a document – one that provides detailed guidance on organizing and archiving unstructured physical materials – has been developed in an organic manner as part of the process of piloting the draft version of this manual. The National Coordination of Widows of Guatemala (CONAVIGUA) was one of those four organizations piloting the draft of this manual that faced the challenge of organizing and archiving their unstructured physical material before they could go ahead with using the manual to create a digital archive. To address this, CONAVIGUA – with support and mentorship from GIJTR – engaged an external archivist to help them organize and archive their physical material. As a result of this process, they created a guide on how to organize a physical archive in 10 steps.

Given that the document has been produced in such an organic manner as part of the project, we have included this guide in the manual in its original form as a direct input from the field from the very CSO that identified the need for it while implementing the draft manual. We therefore point the readers in need of hands-on, practical guidance for archiving and preserving their unorganized physical materials to this guide titled "How to Organize Physical Archives in 10 Steps," developed by Marc Drouin with contributions from Daniel Barcsay and Ludwig Klee, in the Addendum II at the end of the manual.

Cleaning and Backup

Before we start working with the material intended for preservation, we first need to clean and back up the born-digital content.

For our physical items, we should clear a working space and then lay them out, box-by-box, to clean them sufficiently to be handled further. This should always be done wearing protective gloves. At this step, we might note and record any items that might be visibly damaged or degrading.

Anytime we work with digital items, we need to perform an antivirus check to ensure the files are not infected or corrupted. This should always be done by connecting storage media containing the material to a safe computer that is not connected to any computer networks.

Finally, in case you do not have a backup of your born-digital files, you should make one immediately before doing any work archiving them.

Image shared by CONAVIGUA, GIJTR partner organization in Guatemala.

Material in digital form is exposed to a range of risks, from fire hazards through infection or corruption when used in an unsafe computer environment to malicious cyberattacks or simple human error. Therefore, making more than one copy of the digital material is fundamental to achieving a basic level of data security. Further, if our resources allow us to use different types of storage media for backup, we can further lower the risks for our data. Best practices in managing the backup of digital archival material include:

- having multiple independent copies of the digital material
- Copies being geographically separated into different locations
- The copies using different storage technologies
- The copies using a combination of online and offline storage
- Storage actively being monitored to ensure any issues are detected and corrected quickly

At this point in the process, it would be sufficient for two backup copies to be created and stored on two separate storage media, held in two different locations, if possible.

Identification Inventory

The first step in processing the material we wish to preserve is creating an overview of it. In essence, we need to map out what our material contains, in which format, how much material there is, and what state it is in. This should be done on the level of groups of items, not individual documents or objects. Through this process we will create a table containing a list of item groups with key information about each.

The item groups first need to be identified. This is done on the basis of existing information and documentation about the material. Typically, an organization will already have some overviews or lists of different parts of the material. Compiling information from such documents can be a good start, aided by institutional knowledge of the material and any other information we have. This should be complemented by hands-on review of the material, both physical and digital, either by going through boxes and shelves or through review of folders contained in digital storage units. In the process, we should note any

additional or separate groups of items we identify. This will allow us to create the initial list of identified item groups, which we can then place on a table and call "Identification Inventory" or simply "Inventory." In addition to listing the item groups, the Inventory also needs to include information about the type, format, size, amount, condition, location, and storage space/media. An example is provided in Figure 3.

Item groups	Subject	Type	Format	Quantity and/or Size	Condition	Location
Witness statements	Enforced disappearances	Statements	Signed paper copies of statements and audio recordings of statement taking	320 paper statements 100 CDs and DVDs 2 external Hard Disks (1,000 TB)	Boxes marked "ED12" and "ED 13" contain material in poor condition	15 binders in the storage area

Figure 3. Part of an Inventory Table with rows listing the item groups and the columns containing the attributes on which the item groups are described

These are the basic attributes of our item groups that we will need to know before proceeding further with the process of selection, organization, and describing the material. This is also necessary information for development of the General Plan. A brief explanation of what should be considered in assessing item groups on each of these attributes is provided in Figure 4.

ATRIBUTES	DECRIPTION OF ATTRIBUTES
Subject	Identify the subject of the material. What is it about?
Format	Format assessment of a group of items should both include information on their material characteristics (i.e., are the items physical or digital?) and specify their form (i.e., text, photographs, pieces of clothes, microfilms, etc.). Hence, in terms of its format, an item group could be described as "digital documents," "photographs on paper," etc.
Type	Under this category, we answer the question, What are these items? What types of documents, photographs, or objects are they? For example, document types include personal or official letters, reports on a project or an event, filled-in questionnaires, interviews, etc. Photographic items can include, for example, crime scene photographs or photographs of victims, photographs from a workshop, etc.
Quantity	Here we need to note the number of physical objects or digital files contained in the given item group. In case it is not possible at this point to determine the exact number of items, an approximate number should be entered.
Size	This attribute needs to be assessed for digital item groups only by noting the size the given item group takes up on a digital disk.
Condition	This attribute is more relevant for physical item groups, as they need to be assessed in terms of their overall condition and any observable damage or degradation. In the table, the physical item groups can then be categorized as being in "good condition," "damaged/degrading," "poor condition," or "requiring urgent preservation action."
Location	With this attribute, we provide a reference to the location of each item group. For this we need to mark with numbers all physical and digital units in which the material is currently stored. This can include boxes or shelves with physical material, as well as hard disks or DVDs with digital item groups. Once each of these "material containers" is marked with a number, that number becomes the mark of the location of all item groups stored in that particular physical or digital place.

Figure 4. Description of attributes on which item groups are assessed in the Inventory Table

Image shared by CCJ, GIJTR partner organization in Colombia.

Selection and Prioritization

Once the Identification Inventory provides us with a clear overview of what source material we have, how much of it, and in which shapes and forms, we can proceed to decide which groups of materials should be preserved, for how long, and what the order of their preservation should be.

Here, it is important to stress that, given the nature of the material the CSOs documenting mass human rights violations are working with, organizations often wish to preserve everything, as all the material they have collected seems important and valuable. In some situations, this might indeed be the case. However, more often than not, preserving all the source material is neither necessary, reasonable, nor sustainable. For example, a careful assessment might uncover that some of the material is already preserved in another archive, the material does not hold any added value, it came from a compromised source, etc. Further, it could also be due, for example, to the amount or size of the source material or that long-term preservation of it all might be simply unsustainable, as preservation costs might be too high or organizational or technical capacities might not allow it.

Therefore, we need to conduct evaluation, selection, and prioritization of the source material for archiving and preservation. In archival terms, this process is often referred to as "appraisal."

The key step in this activity is development of a set of criteria to which we will evaluate the identified item groups and base our selection and prioritization decisions. These criteria should in turn rely on our Guiding Principles as well as on feasibility, sustainability, security, accessibility, and legal liability considerations.

Selection

Again, there is no universal set of archival criteria for selection and prioritization of source material for preservation. Different types of archives, source materials for preservation, and community and user contexts will affect which criteria are relevant to include as part of the selection process. Figure 5 offers a list of questions that can serve as the basis for devising a specific set of selection criteria tailored to a given archive's characteristics and context.

General	Security and Access	Resources
Is this item group a meaningful part of the archive considering its purpose and its themes? Does its preservation add value to the archive? Does it contribute to achieving the purpose, aims, and goals of the archive? Will the archives' users or a wider community benefit from the item group's preservation?	Are the records in the item group sensitive or confidential or do they contain copyright or other legal restrictions? Can the organization ensure the required data security measures? Does this require additional resources, or added security, legal or other measures? Can access to this material be safely provided to users? What are the ethical and legal risks involved?	Will the size, number, or format of the item group require additional financial, technological, or other resources? If yes, are they feasible? For any additional resource requirements an item group requires, consider whether such requirements are feasible immediately and sustainable in the long term?

Figure 5. Basic set of questions for selection of source material for preservation

Responses to the questions posed in the selection process should be recorded in a Selection Report and preserved for future reference, as decisions regarding the selection and preservation of archival objects will need to be reviewed in later stages and further iterations of the digital archiving cycle. Ideally, during the process, each item group would be listed and responses to the relevant selection questions written down. For the Selection Report, it is sufficient to note the key decisions made in the process and the reasoning they were based on.

We might conclude that certain item groups contain subgroups of items that should be included in the archive, as well as others that should not. In that case, we should divide this item group and separate items that should be included from those to be rejected, creating two or more new item groups, as appropriate. These changes should then be reflected in the Identification Inventory.

As a result of the selection process, each item group should be marked as either "selected for inclusion" or "rejected," in which case it can be removed from the Inventory and the rest of the source material. Alternatively, you might also introduce a category of item groups selected for "potential inclusion," if appropriate. An additional column should then be created in the Identification Inventory and each item should be marked based on its assessment in the selection process.

Documents before organizing the archive, image shared by CONAVIGUA, GIJTR partner organization in Guatemala

Prioritization

The material selected for preservation cannot be processed and archived all at once due to resource, capacity, and technology limitations. Further, some segments of the material might require immediate preservation or digitization. Therefore, it is useful to assess and categorize the level of priority of different item groups for preservation and digitization. In this way, the material in most urgent need of preservation can be given priority. Further, this allows us to plan for any specific security, access-related, technological, or other arrangements that might be needed for the prioritized material.

The main considerations in decision-making on prioritization include:

A. **current state of preservation (i.e., whether the material is damaged or in poor state or could easily be lost or destroyed)**

B. **whether there is an urgent need for availability of the specific material (e.g., for judicial or other transitional justice purposes, or in order to provide important information to the public or key stakeholders, etc.)**

C. **whether to give priority to the preservation of objects that have particular value for the archive, community ,or organization in line with the Guiding Principles**

Specific considerations for prioritization will, however, always depend on the characteristics and context of any given archive, the purpose of the archive, and

TIP!

Handling Material That Is in Poor Condition

In cases where we determine that the source material contains items or item groups that are in very poor condition and need urgent attention, we should immediately but carefully separate and contain such material. This could be the case for both physical (e.g., paper that has mold on it) and digital items (e.g., corrupted or infected files). The physical items determined to be in very poor condition should be put, as soon as possible, in the best possible condition for their further preservation, while any existing damage to them should be contained.

its goals, size, content, etc. Therefore, there might and should be other tailored criteria of prioritization developed for any given archive.

Resource Recommendation!

Listed here are useful resources detailing the procedure for caring for different types of materials (paper, photograph, audio, video, etc.) in poor condition.

"Preservation and Conservation: Caring for Personal Collections" from Emory University Libraries

"Disaster Response and Recovery" from National Archives

The outcome of the prioritization assessment should be the classification of each group of the material selected for preservation into priority classes (e.g., Priority levels 1, 2, and 3). Accordingly, an additional column should be added to the Identification Inventory wherein each item group will be marked in line with its assigned level of priority.

6.3 Organization and Description

Once we have an Inventory with basic information about the selected item groups for preservation, we can proceed to organize and describe the material. This is a necessary action to allow further archival processing and preservation, as well as to ensure that the future archive is structured, which will allow it to be manageable and searchable and ultimately that its content is accessible. This step has major importance in the process, as it will be the basis for the structure of our future digital archive, with repercussions on all aspects of its development.

Organization

Organization of material for digital preservation involves introducing a certain logical and hierarchical order into it and thereby devising its structure. This is done on the level of item groups identified through the Inventory, using an organization's knowledge and understanding of the material.

The process of organization of the selected material entails the entire content of selected material being divided into several fundamental groups, each based on one or more common features shared by item groups they contain. These most-generic groups are then divided into smaller subgroups of material, and so on down to the level of individual items.

The most generic groups of material are often referred to as "collections," or in strictly archival terms "fonds." Each collection is divided into "series," which can contain individual items as well as "subseries" and "folders" (sometimes also referred to as "files"), which are smaller, subordinated units of structure that then also contain individual items. See Figure 6.

Resource Recommendation!

Here, we recommend you consult a detailed and highly instructive description of the process that can fruitfully be applied to organize unstructured sets of material provided in a recent GIJTR publication

"Living Archives – An Introductory Toolkit for Civil Society Organizations in the Creation of Human Rights Oral Archives and

Organizing Their Documentation."

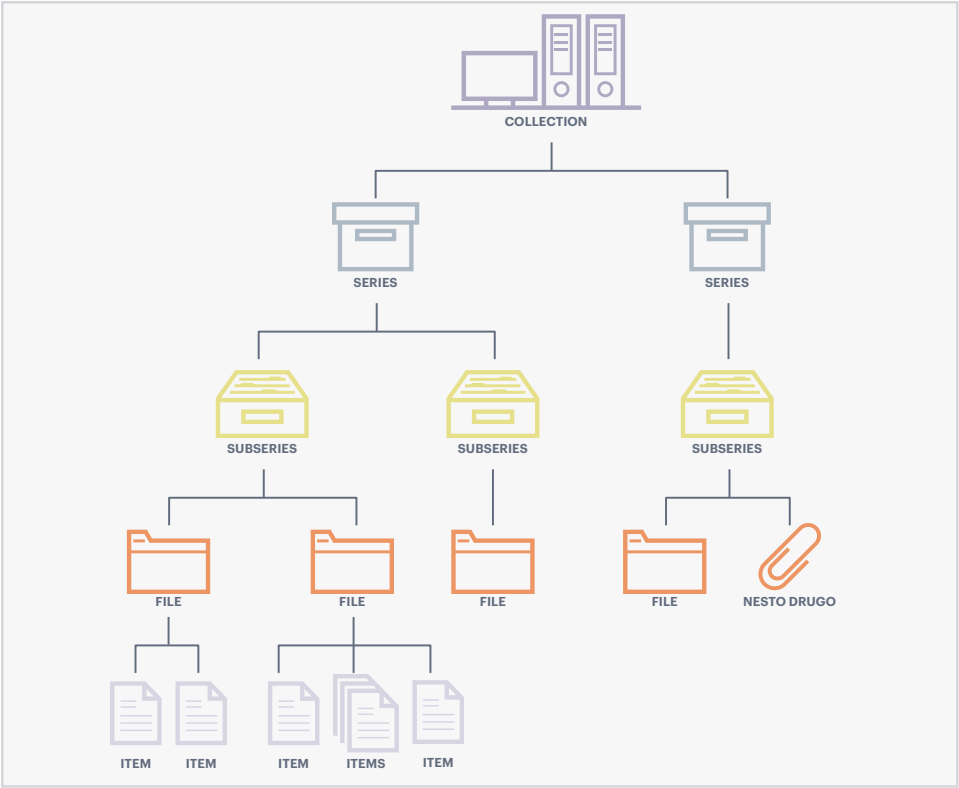


Figure 6. Diagram showing an archive's structure

This process of grouping, ordering, and devising the structure of the material – which, in archival terms, is referred to as "arrangement" – cannot be conducted by following a cookbook-like instruction manual. It requires analysis and consideration of the material and the context in which it was created, discovered, or received. The goal is to devise a structure and order that will preserve as much of the original context of the material as possible, including the information and meaning contained in the original relationships between groups of material. To achieve this, the archival rule of thumb is to arrange the material with respect to its "provenance" (i.e., origin or creator) and "original order." This means mirroring, or following to the greatest extent possible, the structure and order that is already contained in the material itself. The presumption here is that there is either an obvious or underlying logic and order to the organization of any given group of material selected for archiving, and that in the process of organizing the source material we can identify or uncover this logic and then mirror it.

However, this approach is applicable only in cases where there is a clear or discernible order and structure to the material. This is often not the case for CSOs aiming to create digital archives of various material related to human rights violations. On the contrary, while some segments of the material CSOs are working with might be structured and ordered, there will usually be larger sections of it that are only partly or inconsistently structured, or that have no order to them at all.

In such cases, we should not attempt to preserve the "original chaos" found in the material. Rather, we should proceed to arrange it in a way that will best facilitate its use and management, while also relying on the analysis of the material itself. This can be done by devising several possible criteria for grouping of the material (e.g., based on its authorship, the function it served, the action it was a part of, or similar). These initial criteria can then be tested by applying them to a sample of the material. Based on the feedback from this piloting process, through which we will eventually identify the criteria that best suit the material at hand and allow for developing a distinctive structure of collections and series into which all item groups can be logically placed, we can tweak the test criteria further.

The result of this exercise will be an archival structure of collections, series, and, if appropriate, subseries and folders into which all identified item groups can be

Image shared by CONAVIGUA, GIJTR partner organization in Guatemala.

logically and meaningfully added. This archival structure can then be visually represented through a hierarchy tree or similar scheme.

To complete this stage, we need to revise our Identification Inventory and turn it into a table that reflects the newly developed archival structure so that its elements – collections, series, subseries, and folders – become the main unit of analysis. An example is provided in Figure 7.

This table of archive's structure will be a necessary tool for the next steps in archival processing of the source material, as well as for further development of the General Plan and completion of the Planning and Organizing stage.

Collection	Subject	Type	Format	Quantity and/or Size	Condition	Location
Collection 1						
Series 1						
Series 2						
Subseries 1						
Subseries 2						
Series 3						

Figure 7. Table of archive's structure with collections, series, and subseries as units of analysis

Description

Now that we have organized our archival material, we need to describe its content in a way that will allow anyone to search for, locate, and access items in the collection. Description of archival material also enables its proper preservation and guides future users by providing important contextual information. It further enables establishing connections between items, even from different series. Simply put, without description, an archive would be more of a form of storage in which it would eventually become impossible to find or manage content.

The first decision an organization needs to make at this point is whether the archive will need – and whether it is even possible – to include a description of each individual item or instead just describe the content according to the level of item groups (i.e., folders or above).

A basic level of description of the material on an item level, at a minimum by identifying each item with a unique identification number, is necessary for further processing. For born-digital material, this can be easily done using software (which will be discussed later), while for physical material, we need to go through each individual item manually and identify it. More detailed description of content on the item level is certainly preferable, as it allows for it to be more easily searched for and located, as well as because it provides more detail and context, which all significantly improve future preservation and access. However, this might not always be possible, despite those potential benefits being of essential concern for human rights archives.

The source material could contain an extremely large number of items, thus making it impossible to describe each one; or the organization could be unable to raise the needed funds and resources; or there could be a lack of time due to urgency to proceed quickly for safety or preservation reasons. Whatever the case, a decision should be made by carefully weighing the benefits on the one side (in terms of improved access and preservation) and the downsides on the other (including feasibility, time, and resources required).

The same trade-off applies to the second main decision we need to make at this stage, which is in how much detail and with how many elements do we want to describe our content, being mindful of our practical limitations. Including more elements of description will allow us to provide better access and more contextual guidance to users, but it will also require more time and resources. Again, each organization needs to select the elements of description based on the individual circumstances: size and characteristics of its archival material, type of access it needs to provide, and organizational capacities.

The elements of archival description, also known as "descriptors," provide data on location in the archive's structure, physical and technical characteristics, informational content, and the function or purpose of archival item(s). There are different groups of such descriptors, the most relevant being general, content, and technical descriptors.

General descriptors record information that identifies and locates items, folders, subseries, or series within the archive, such as:

- A unique code or number
- Series/subseries/folder
- Title, author, date of creation

Content descriptors record information contained in an item, folder, subseries, or series regarding such categories as:

- Theme
- Location
- Time
- Actors

Technical descriptors record physical and technical characteristics of an item, folder, subseries, or series, such as:

- Storage location
- Storage media
- State of preservation
- Format, volume

In devising these descriptions, as well as in selecting and formulating the descriptors for our archive, we should rely on the widely accepted and used archival standards and instruments that can assist us in this process. The recommended standard for archival description is "ISAD(G)," developed by the International Council on Archives (ICA), which also has its application devised specifically for human rights archives. There are also other relevant standards that could be used.

There are also numerous other possible descriptors, some widely used, others specific for a given archive. Each archive will select those best tailored to the needs of its material, bearing in mind its Guiding Principles.

For CSO human rights archives, a particularly important set of descriptors is one recording private, sensitive, or confidential information present in the archival content. It is essential for a human rights archive to be aware of any such legally or otherwise protected material so that it can appropriately manage and control access to it.

Our task at this point is to analyze and review the content and its context in order to describe at the level of description we selected – items, folders, subseries, or series – in relation to each of the descriptors selected. We then record these descriptions in the table of the archive's structure we created in the previous step.

With this, we have completed the organizing phase of archival processing of our invaluable material for preservation. Now we must leave the safe haven of established and standardized archival procedures and depart into the restless seas of digital archiving, where ever-changing technologies are, for better or worse, inseparably intertwined with archival processes.

The first step on this journey is selection of the software and hardware framework of our future digital archive – the Digital Archiving System.

6.4 Digital Archiving System

A Digital Archiving System is a technological infrastructure of a digital archive. It defines the scope and limit of the archive's functions and is instrumental in the archive achieving its primary aim and goals and upholding its General Principles. Therefore, selection of a Digital Archiving System has to be built into the planning stage of digital archive development as its essential element.

Image shared by CCJ, GIJTR partner organization in Colombia.

What Is a Digital Archiving System?

The main goal of digital archiving is to ensure that the invaluable content we are preserving remains unchanged and accessible long into the future. This can be achieved by implementing an adequate and sustainable technological framework, or as the Digital Archiving System, for our digital archive.

A Digital Archiving System is a system of software and hardware components that consists of databases, software tools that manage databases, and storage media. A database stores information about the archive and its contents in an organized collection. Any table, such as the Inventory we created in the previous step, could be seen as a form of a rudimentary database containing information about an archive. An archival software tool then allows for management of a series of such databases, their content, and relationships between them. Archival software also serves as an interface between databases contained in the Digital Archiving System and the system's users. The software enables us in practice, for example, to add item groups to our Inventory or create a new subseries.

The databases and software tools are merged together and make up the main software component of the system: a digital archiving software that allows us to manage an organized collection of information about the archival material.

The digital archival material itself, however, is located on storage media – typically physical devices that can store, retain, and make digital archival data available for retrieval. Well-known examples of storage media include a hard disk drive, flash memory, and DVDs. Until recently, digital content was stored only on individual pieces of different types of storage media, such as a single hard disk or CD. However, in the past two decades, two new forms of storage media have emerged: server-based storage systems and cloud-based storage.

A server-based storage system is usually located at the archive's premises and comprises multiple storage media contained within a server that provides additional protection and allows for the recovery of data in case of a failure. Their set-up and management require advanced IT expertise.

Cloud-based storage is, in essence, outsourced server-based storage – a commercial service providing online storage and access to data. It is relevant to understand that when we store our data in a so-called "cloud," it is in fact stored on a large-scale, server-based system of a company we hire for this service.

Functions of the Digital Archiving System

The software and hardware components of a Digital Archiving System work together to enable performance of the key functions of a digital archive, including storing, backing up, preserving, maintaining integrity and authenticity, safeguarding, providing access, managing, and eventually migrating archival data. In support, and to complement these main functions, a Digital Archiving System needs to enable us to perform a whole range of specific tasks and actions (e.g., checking data for errors, restoring lost data from backup, restricting access to sensitive data, and many others).

Given that a Digital Archiving System performs such an essential role, it is critical to select software and hardware solutions for it that will adequately provide for the specific needs of a given archive, which define the requirements it will have from a Digital Archiving System.

Such requirements are always specific for any given archive. In defining them, we should relate them to the General Principles of the archive: its purpose, aim, goals, and responsibilities. We also need to consider practical, logistical, and resource-related aspects of the Digital Archiving System we choose to implement, as well as our organization's current and potential capacities to support it. Figure 8 lists some of the aspects of our archive that require consideration and analysis when selecting a Digital Archiving System.

Content and Standards	Admin and Resources	Access, Use, Security
Type and format of the material to be standards (e.g., for description, preservation, data security, etc.) need to be applied Volume of material that needs to be preserved How long the material needs to be preserved in the digital archive Standards – e.g. for description, preservation, data security etc. – need to be applied	How the digital archive will need to be managed and administered Whether we need to record the chain of custody over the material Financial, technical, human, and other resources needed for acquisition of digital archiving software, storage media, and maintenance	Modes in which the material from the digital archive will be accessed and used, and by which user groups The intended uses of the archival material Specific concerns regarding data security

Figure 8. Considerations in selection of a digital archival system

Responding in writing to these and other points that might be relevant for a concrete archive will give us an overview of its specific needs. These can then be translated into our main requirements for a Digital Archiving System. In selecting components for our Digital Archiving System, we will seek solutions that, as much as possible, meet these requirements.

The process of selection of the Digital Archiving System software and hardware components should be recorded and documented in terms of analysis and reasoning on which it is based. Documenting the process facilitates any future modifications, upgrades, and eventual migration of data to new Digital Archiving Systems.

We provide example lists of main requirements for a digital archiving software and storage media for a CSO digital archive in the Addendum III. Again, each archive will need to include its own tailored lists of requirements. Moreover, these main requirements will need to be further devised and specified as the selection process progresses and concrete software and storage media solutions are being reviewed and considered.

Selecting Digital Archiving Software

The archival software solutions come in different shapes and sizes and there is a wide range of options available. Different options provide for a different set of functions and vary in quality of performance. They also differ in terms of financial and human resources, technical expertise, and organizational capacities their purchase, implementation, maintenance, and development require.

The key distinction is between commercial digital archiving solutions sold by software companies and open source software developed by communities of programmers that, importantly, is free to use. Both these options have their benefits and their disadvantages that need to be carefully considered before we make the selection. The list of requirements from the Digital Archiving System (see Addendum III) will provide useful guidance in this process, as the two different types of software solutions can be evaluated against it.

The key distinction between open source and commercial digital archiving software is not whether one is free and the other is not; rather, they are grounded in differing methodologies, approaches, and sustainability models,

Documents before organizing the archive, image shared by FAMDEGUA, GIJTR partner organization in Guatemala

Image shared by CCJ, GIJTR partner organization in Colombia.

which leads to them having advantages in some areas and disadvantages in others.

In a nutshell, by selecting one type of software over the other, we select between prioritizing either flexibility or usability of our archive. Open source software is more flexible and allows for quicker and innovative changes to the structure, elements, and functions of the archive. At the same time, however, it requires more time, effort, and expertise to use, maintain, and develop than does a commercial solution.

Another essential dilemma in this process is whether we will select a comprehensive, all-in-one solution or a modular solution, which is a combination of individual software tools working together in one system. The former provides all archival functions within one software solution and is usually more user friendly for management and use. The latter provides more opportunities to fine-tune the system's functions and introduce new options or services.

Once we have made the strategic decisions on whether we will implement an open source or commercial digital archiving software, and an all-in-one or a modular solution, we should select a concrete product among the many

available. Our list of requirements will again serve to identify the products that provide the best possible fit for our digital archive in terms of functions, actions, and tasks it is required to perform while being feasible and sustainable in light of required resources.

It is a good practice to test, on a sample of material, several software solutions that you are being considered in order to test their compatibility with our archive and get a better sense of their look and feel, functionality, and efficiency.

It is not advisable to apply software solutions that have recently been developed and therefore not yet been widely applied and tested. Rather, we should opt for a proven and widely used solution and carefully analyze available information on its performance, evaluations, and user experiences. Reaching out and directly exchanging experiences with other CSOs that are considering or implementing such software solutions would be particularly beneficial.

Selecting Storage and Backup Media

Similar to the selection of digital archiving software, we should make a decision between the main types of archival storage media and their respective advantages and downsides. The most frequently used storage and backup media for archiving include external hard disks (e.g., HDD, RAID, SSD, or flash storage), optical disks (e.g., CD, CD-ROM, DVD or Blu-ray), magnetic tape, server-based storage systems, and online cloud storage.

When selecting storage media, the solution might involve more than one type of a product, as such a strategy would improve data safety and backup. For example, if resources allow it, we could opt for HDD external hard disks as the main storage media and use online cloud storage as a backup.

Different archives will have different priorities in setting the selection criteria. However, there is a set of dimensions that are almost universally considered relevant, including storage media longevity, capacity, viability, obsolesce, cost, and susceptibility. A useful overview of these criteria, along with other information relevant to the storage media selection process is provided in the UK National Archives publication "Selecting Storage Media for Long-Term Preservation."

PART 7

3:00

DIGITIZATION, PRESERVATION, AND INGEST

Now that we have completed the planning and organizing stage and come out the other side safely armed with the General Plan, the table of the archive's structure, descriptions of the material, and a decision on software and storage media for the Digital Archiving System, we are prepared for the next stage. This is where the actual magic happens: the creation of our digital archive.

Along with the great promise it brings, this stage is also the most dynamic and complex, as well as the most resource-heavy, expertise-driven, and technologically demanding for the organization.

Our goal at this stage is to process and prepare all selected material – both physical and born-digital – and to make it digital preservation-ready. This means that by the end of this stage we will have the material prepared with respect to all necessary technical and archival requirements for transfer into our newly selected Digital Archiving System. This includes a series of actions using software and other technological tools that need to be applied to our selected source material to be able to properly archive it and preserve it long term.

Additionally, if we are working to digitally preserve source material that is wholly or in part physical, then this stage includes a major pre-step: digitization.

7.1 Digitization

Through the process of digitization, we create digital copies, or "surrogates," of original physical items. These digital copies are then processed as digital archival objects, preserved, and made accessible. We will therefore be focusing on preservation of these digital copies rather than the original physical items. Consult Addendum II for further guidance.

There are different types of physical objects we might want to digitize that can be stored on a variety of media. They include, for example, text, photographs, drawings, maps, video, audio, and other types of content stored on paper, audio cassettes, 16 mm tape, or any other physical or analogue storage media.

Image shared by FAMDEGUA, GIJTR partner organization in Guatemala.

They could also include objects such as pieces of clothing, banners, personal belongings, etc.

Clearly, the type of material we need to digitize will define both major and specific decisions to be made in the process – and each organization will make them in line with its goals and capacities. However, there are also general elements of the process that need to be addressed in all digitization projects. This chapter outlines those elements of digitization, which are relevant to the process regardless of the type, content, or storage media of the material.

BREAKING NEWS

In-House Digitization May Cost More Than Outsourcing

If the capabilities of the organization are not sufficient for the requirements of the digitization process, a decision to hire an external company for the project must be considered. Doing so may determine the success or failure of the program. Initiating digitization with inadequate preparation, resources, and capacities could produce more costs than results, with little or no long-term value. On the other hand, a quality-assured, well-planned, and executed outsourcing option could save substantial time and effort. Hence, in-house digitization, with the different costs it involves, may sometimes cost the organization more than outsourcing the work externally.

Digitization is a major, demanding archival project in and of itself, and requires due attention, careful planning, and dedicated implementation. Since we are looking at digitization as part of a larger process of building a digital archive, we have already discussed some of the issues involved, mostly regarding the first few stages of the process. An overview of the digitization process is outlined in Figures 9a and 9b.

1. Planning

General: goal, outcomes, timeframe, resources

Logistical and organizational: workflow, conditions, space"

revise to "naming, equipment, metadata

Archival and technological requirements: quality, format, file naming, equipment & metadata

Planning for preservation of original physical items

2. Preparing Material

Creating inventory of physical material

Review of material and selection of material for digitization

Description of material

Preparing physical items for digitization

Figure 9a. Overview of stages and actions in the digitization process

3. Preparing Data/Tech

Defining digitization requirements, file naming, format selection, standard of quality, collection of metadata

Obtaining and installing digitization equipment, software, storage media

Setting up equipment to meet digitization requirements, testing, fine-tuning

4. Implementation

- Preparation of material
- Process scheduling
- Digitization
- Quality control
- Post-processing and OCR
- Storage and backup

Figure 9b. Overview of stages and actions in the digitization process

In previous chapters, we discussed development of a General Plan, the creation of an inventory, and selection and description of the material – which are also the first steps of the digitization process. Hence, having already covered the first two, we can pick up the digitization process at the beginning of the third stage with the preparation of archival and technological elements of the process.

BREAKING NEWS

Digitization Can Be Done on a Small Scale and With a Modest Budget

Small-scale digitization projects need to be adjusted to fit modest capacities and resources. Generally, that means there may be only one or two persons tasked with performing all the steps of the digitization process, on one computer and with limited resources. The process is certainly less efficient, less reliable, and slower under those conditions, but it is doable and – whenever other options are not available – it is highly recommendable. Any digitizing work you can conduct can be highly significant, especially if the material is fragile and prone to deterioration.

Specifying a Naming Convention for Digitized Files

For a digital file intended for archiving and preservation, a name is not just a name. It is also a very important descriptor of that particular item, which should contain information that allows us to identify what the item is and what it contains so we can locate it in the archive and properly manage and preserve it. Therefore, an important element of specifications for the process of digitization is development and application of a consistent set of rules, a so-called "naming convention" for digital surrogates we create from physical items.

There are no universal rules for file naming, and each organization needs to develop its own naming convention that best suits its archival needs. However, the name of a digital surrogate should always provide a reference, a connection between itself and the physical item from which it was created through digitization. In principle, a file name should contain several components that identify it, for example its unique identifying number, or its date of creation, or a reference to its content, or series, or subseries, or folder it is a part of.

We should also bear in mind that these file names primarily need to be processed and understood by software we will use for managing our digital archive. Hence, our primary concern in naming files is to apply a convention that will enable our Digital Archiving System to correctly identify the file and use information contained within it. However, many consider it a good practice to also include a

descriptive component in a file name, one that could be understood by humans as well, for example a reference to its title or content.

While, as mentioned, there are no strict instructions for developing a naming convention, we can nevertheless identify some basic recommendations, as outlined in Figure 10.

General	Identifiers	Standards
Use a reasonable number of components for a file name. Names should be as short as possible, so use abbreviations. Be consistent in application of the file naming convention and do not allow for exceptions.	Include key identifiers as components of a filename (i.e., identifying number of the item). Include descriptive components such as date, title, or reference to its content.	Use only English alphabet letters (a–z), numbers (0–9), dash (-), and underscore (_). Dates should be entered in the ISO standard format (i.e., yyyy-dd-mm).

Figure 10. Recommendations for a file naming convention

Specifying File Formats and Quality

In addition to the file name of a digital surrogate, its digital format and the standard of quality to which it will be digitized also need to be specified before the process can begin in earnest.

Given that the same type of files – such as documents, photographs, or video – can be stored in different digital formats, we need to specify which formats we will be using for the digital surrogates created from our physical items.

Given that we are digitizing material with the purpose of long-term preservation, it is important that we select formats that will allow their proper viewing and use in the future by new generations of software. To prevent our digitized files from becoming obsolete, we should choose formats that are robust and resilient to change over time.

This means we should be looking for formats that meet the necessary standards, which are well-established and widely used with substantial and positive user

feedback. The formats we select should also allow us to add information and metadata to the files and should have stable support, either commercially or through an open source community.

Clearly, we will be considering different sets of formats depending on the type of items we are digitizing – documents, photographs, video, etc. The scope of format options can be overwhelming and there is no universally ideal solution for each type of digitized content. The selection, again, depends on specific needs and circumstances of the archive. Nevertheless, there are formats that have a proven high level of robustness and resilience to change. Figure 11 provides an overview of such formats for the most frequently digitized types of physical items: documents, pictures, audio, and video.

Physical Item Type	Robust digital File format
Documents	PDF
Photographs	RAW or TIF
Slides and negatives	RAW or TIF
Audio	WAV
Video	MP4

Figure 11. Overview of robust digital formats for digitization of different types of physical items

Specifying Quality Standard(s) for Digitized Files

An important element of the specifications for the digitization process is the standard of quality to which we want and need to digitize our physical items. This is usually referred to as the "resolution" of a digitized document, photograph, or video. A higher resolution of a digital surrogate will allow for a better user experience and wider possibilities for its use – and overall a better copy of its original than a lower-resolution file. However, higher resolution also means that the digital surrogate will have a bigger digital size and will therefore take up more space in our storage media.

Image shared by FAMDEGUA, GIJTR partner organization in Guatemala.

Therefore, in specifying the resolution of the digital surrogates we will create, we need to weigh the requirements for their quality standard with the demand it creates in terms of digital storage space for our archive.

As human rights organizations working with unique and invaluable material, we can easily be tempted to digitize all our material in the highest available resolution to ensure the best possible quality of the digital surrogates. However, this would be neither feasible nor sustainable, as it would create immense difficulties in not only storing, but also processing and preserving such files long-term. Organizations therefore need to make digitization quality specifications in line with their goals and capacities. As a guidance, Table 12 provides an overview of what is often considered to be minimal and optimal resolution quality levels for digitization of different types of physical items.

Item Type	Minimal Quality	Optimal Quality
Documents	300 DPI	600 DPI
Photographs	600 DPI	1,200+ DPI
Slides and negatives	1,200 DPI	2,400+ DPI
Audio	16-bits and 44.1 KHz	24-bits and 96 KHz
video	1080P or 2 Megapixel	2K+ or 4 Megapixel

Figure 12. Overview of minimal and optimal resolution quality levels for digitization of different types of physical items

Metadata: Descriptions of Digitized Files

In the section dealing with the planning and organization of a digital archive, we discussed the important process of describing the archival material on several of its relevant attributes and creating a connection between those descriptions and the material by recording them in a table. This is necessary, as it allows us to later search for, locate, and identify items and item groups based on those descriptions, as well as to properly manage, preserve, and use the archival material. The same principle applies for the digital surrogates.

After digitization, it is the digital files we create from the physical originals that will become the items in our digital archive. Hence, they also need to be described and have their descriptions attached to them so they can later be found, accessed, and preserved.

These linked descriptions of archival items are known as “metadata,” or data about data.

In the process of digitization, it is essential that relevant metadata is collected and attached to the digital surrogates we create. This is because, without its attached metadata, a digital surrogate becomes meaningless and unusable – as we might be unable to find or identify it or understand what it is, its context, history, creator, or where it belongs in the archive.

Most of the metadata we need to preserve is linked to the digital archival files they describe and created and captured through software tools we use to digitize, manage, and archive the data. This includes basic metadata (e.g., date of creation/digitization) as well as very technical types of metadata, such as those on validity or integrity of digital files. The software tools can therefore allow us to capture the metadata. Concrete technical solutions in relation to different types of metadata being captured and preserved are discussed further in the manual. However, our main concern here is to select which types of metadata we want and need to record and preserve in our digital archival files.

Compared with physical originals, digital surrogates require, and allow for, a whole range of additional metadata to be collected. This includes metadata such as technical specifications of an archival digital file, as well as information about its creation and any further digital action taken on it. For CSOs working with human rights material, such technical metadata is important not only for preservation but also for maintaining credibility of a digital surrogate and establishing the chain of custody.

There is a wide variety of types of metadata that could be collected about digital surrogates both during and after the digitization process. The most common types, based on their purpose and function, are summarized in Figure 13.

Descriptive & Structural	Admin & Preservation	Technical
Descriptive metadata gives details about a digital record and its content to make it easier to find. **Structural metadata** provides information about the internal structure of a digital file, including information like page, section, or index.	**Administrative metadata** refers to the information about the management of a digital record, such as who created it or who can access it **Preservation metadata** helps usage of digital records in the future; includes information about what software or hardware is needed to open and use a digital file	**Technical metadata,** rather than being created for the purposes of archiving, is often captured automatically through the software or hardware used to create a digital record. For example, photos created by a digital camera automatically capture information about the image and embed this information in the file itself.

Figure 13. Types of Metadata

Selecting the metadata for any given digitization project will depend on its context and circumstances: an organization's resources and capacities, the type of material, intended applications of it, types of access and user needs, among others.

Existing metadata standards and specific, tested, and widely used metadata profiles and sets provide guidance through the maze of numerous metadata types and formats. However, there are now so many different metadata standards and sets developed and proposed by different organizations that their sheer number by itself creates an obstacle to identifying those we want and need to use.

A good place to start is with the so-called "Dublin Core Metadata Element Set." Dublin Core is a very widely applied set of 15 properties or elements for describing digital files. These elements are often considered to be a standard set of metadata that are applied almost regardless of the type of archival material, theme of the archive, or type of software used in the Digital Archiving System. Further, for preservation purposes, the so called PREMIS metadata standard provides a useful reference and guidance (PREMIS: Preservation Metadata Maintenance Activity (Library of Congress)).

Whatever set of metadata we select for our collection, there is another set of decisions that we need to make about them to complete their digitization specifications. These include questions such as, Where will the metadata be stored? How will it be captured? When in the process do we capture it?

Making considered decisions related to these questions in advance of the digitization process will provide us with a plan for standardized and consistent collection and structuring of metadata throughout the digitization process. This is important in order to make our metadata "interoperable" – which means making it structured and formatted in a way that allows it to be read and used by different computer systems.

Making our metadata interoperable will save us significant time and resources (as well as headaches) later in the process, not least in the next step when we need to ingest and make operable that metadata, along with the digital surrogate files to which it is linked, in our Digital Archiving System. These issues related to the processing of digital files and their metadata will be discussed in more detail in the upcoming section, where we look at how our entire material – both digitized and born-digital – needs to be prepared for ingest into our Digital Archiving System.

Selection, Set-Up and Testing of Equipment: Software, Hardware, and Storage Media

This manual cannot provide recommendations on specific digitization equipment, software, or storage media, or how it should be set up and optimized. Such advice would necessarily be too generic for requirements of any concrete project, and it would also be likely to quickly become obsolete.

However, we should mention three elements that need to guide our decisions in selecting the technology we use for digitization: characteristics of the material, an organization's capacities and resources, and the archive's needs and requirements.

First, the equipment we select and how it will be set up and fine-tuned depends on the material we are digitizing: type, format, state of preservation, size/length of the originals, and quantity. Fragile material, for example, will require more refined and sensitive equipment and set-up, while large quantities of material will urge a solution that provides for quick processing.

Image shared by CONAVIGUA, GIJTR partner organization in Guatemala.

Further, our decisions will be dictated by the resources we have in terms of time, expertise, staff, space, and finances. Each of these aspects will set limits on what can be a feasible solution for our project.

BREAKING NEWS

More Expensive Equipment Can Bring Down Overall Digitization Costs

We should be mindful that although digitization can be done on a different range of budgets, it is important to look at total costs of a project rather than one-off costs separately, such as the cost of a piece of equipment. Total project costs should include staff wages, equipment, time, etc. More expensive equipment that processes items more quickly, for example, could save us much more than it costs if we also calculate staff time and wages.

Finally, and most importantly, the needs of our archive and its future users as well as the modes of planned use for the materials we are digitizing should define the minimal and optimal requirements of the equipment.

For hardware and software, regardless of the type of material (documents, photographs, video, or other), the requirement will be to provide digital surrogates of desired quality in adequate formats and capture the selected metadata. In terms of storage media, the most important aspects to be considered are its reliability (resilience to data loss), durability (usability over a longer time period), and scalability (potential to expand the data storage space as required).

Once we have selected and obtained our equipment, we need to properly install and set it up in line with our digitization requirements. This process is important and needs to be done properly, otherwise even the right equipment will not yield required results. Hence, if an organization does not have internal expertise, external assistance would be advisable at this point.

This especially true given that the set-up and its fine-tuning are not a one-off activity, as the process requires repeated testing and iterative changes before the required result is achieved. The testing process should include a sample of different groups of materials and involve the entire process of an item's digitization (i.e., the digitization workflow).

Image shared by CONAVIGUA, GIJTR partner organization in Guatemala.

Implementation: Digitization Workflow

The final stage of digitization is implementation of all the different elements that we have been planning, deciding on, and devising in the previous stages. Digitization is a complex process, but if all of its parts and functions are planned and designed well and in advance, its implementation will be streamlined and fruitful.

That is why in putting all elements together we should develop a detailed digitization workflow, which should include all its actions and operations – starting with review and preparation of physical items and workspace to the completion of the workflow through storing the created digital surrogates and making backup copies.

Each digitization project will have its own unique workflow and specific sequence of digitization actions and operations. Further, some activities, such as quality control, will be repeated at different stages of the process, while others will be executed simultaneously or in parallel. Although specific actions and their sequence are tailored to each concrete project, we can identify the key elements required in any digitization workflow: preparations, process scheduling, digitization, quality control, post-processing, and storage and backup.

Preparation of Material, Protocols, and Workspace

The digitization process begins in earnest with ensuring a clean and appropriate workspace, allowing enough area for work with physical materials as well as for the digitizing equipment and a computer. Assuming that fragile or otherwise compromised material has already been removed, we can proceed to cleaning our physical material and removing any added items, such as paper clips or staples on documents.

Information and relevant digitization specifications about file naming, file resolution, and format, plus any metadata to be recorded should be on hand and well-organized.

Process Scheduling

As part of the workflow, it is essential to clearly schedule the entire process – that is, to determine, document, and then strictly apply an exact sequence of operations to be performed during the digitization process. The scheduling should include buffer time for unexpected events.

Resource Recommendation!

Excellent examples of digitization workflows and scheduling for organizations dealing with the preservation of cultural heritage material are provided in "Technical Guidelines for Digitizing Cultural Heritage Materials," issued by the USA Federal Agencies Digital Guidelines Initiative.

Digitization Processing

The process of digitization itself will clearly be very different depending on the type, volume, content, and other characteristics of the material. Paper documents and photographs can be scanned reasonably quickly, while analogue audio and video will need to be digitized in real time. Artwork and historical documents will require a different scanning specifications set-up than will an administrative document.

Regardless of the differences, a good practice at the start of each digitization session is to digitize a reference item (document, photograph, short sample audio or video) with the result reviewed against specifications as a form of ad hoc quality control. In case of any discrepancy from the digitization specifications, equipment can be checked and its set-up fine-tuned. This will help avoid wasting entire sessions of work due to equipment or set-up issues.

Post-processing

Post-digitization processing of digital surrogates includes making slight corrections to a file to adjust it to a certain standard or specific project specification. This could include actions such as increasing the sharpness of sound in a video file or brightness of an image on a document.

Post-processing might sometimes also include creation of secondary, derivative copies of the file. These are created for specific purposes such as providing access or producing high-quality reproductions, and also for creating fully searchable documents from originally non-searchable image files through the application of Optical Text Recognition (OCR) software. In essence, by running OCR software on our scanned image of a document, we add a layer of text onto that image file so other software can read it, which makes the document fully searchable. This is essential for making human rights archives more accessible and visible, which is often a key purpose of their digitization. Given the importance of the application of OCR technology in creating fully searchable text files from our digital surrogate image files, in Addendum IV we provide a set of recommendations regarding its use.

Quality Review

There are two elements to digitization quality control, and both can and should be implemented at multiple points in the process scheduling (i.e., both during and after digitization, as well as at regular intervals over the course of the project). The first element relates to ensuring that all physical items intended

Digital Forensics

If working with older data storage formats or digital material of unclear origin and features – especially when a history of the material and "chain of custody" need to be established – a promising area of development is "digital forensics," which provide benefits in addressing digital authenticity, accountability, and accessibility. This forensic technology can make it possible to identify privacy issues, establish a chain of custody for provenance, employ write protection for capture and transfer, and detect forgery or manipulation. It can also extract and mine relevant metadata and content, enable efficient indexing and searching by curators, and facilitate audit control and granular access privileges. Digital forensic technologies vary greatly in their capability, cost, and complexity, with certain equipment ranging from free to expensive. Some techniques are very straightforward to use, while others have to be applied with great care and sophistication. There is an increasingly rich set of open source forensic tools (e.g., "BitCurator") that are free to obtain and use.

for digitization have indeed been digitized. This can be done automatically by comparing the two sets of data for physical items and their surrogates; however, this should also be accompanied by a sample manual check to ensure that digital surrogates properly correspond to their physical originals.

The second element of quality review is ensuring that the digitization specifications have all been met – that the digital surrogates are created in the right format and quality, with correct filenames, and selected metadata has been captured. Here again we will need to use a combination of manual and automated quality review, which is supported by software tools and applications such as "JHOVE."

Storing Digitization Products

At the end of the process, we need to temporarily store the products of digitization on one or more storage media until they are prepared and ingested into a digital archival system. The end result of the process should be one or more digital surrogates of the original, which are often referred to as "master files." These are stored in a file directory structure created for this purpose.

Master files are the best-quality files we produce through digitization and are intended to be preserved long term without loss of any essential features. The number of master files we will create will depend on the content of the originals and the planned uses of the digital surrogate.

In addition to master files, we can also produce a number of secondary files, often called "access" or "service files." These files are created from the master file and optimized for the intended use (e.g., for web or for research).

For organizations working with documentation on human rights abuses, it is especially important to note that these derivative files are used for the creation of files with fully searchable textual content through OCR. The usual practice is for only master files to be stored for preservation purposes. However, given the importance of the OCR – and therefore fully searchable versions of documents – for human rights archives, it is advisable to also create and store two such readable files, one as an access copy and the other for preservation purposes. The same applies for the master files, as we should create at least two backup copies and store them on two separate storage media whenever possible.

Image shared by ASOMOVIDINQ, GIJTR partner organization in Guatemala.

7.2 Preservation and Preparation for Ingest

We are now fully in the digital archival world.

All our material is now in a digital form.

We also have a digital archival repository – in the form of a Digital Archiving System.

To complete the process of creating a digital archive, we now need to employ a set of software-based digital archiving techniques on both our digitized and born-digital material. This is necessary to prepare it for ingest and long-term preservation in the Digital Archiving System. We also need to set up and prepare our Digital Archiving System itself – its databases and software tools and applications – to properly receive, store, and preserve our digital archival material.

To do that, we first need to review our basic archiving tools – the archival structure table and descriptions of material – which in this digital archiving world will take the form of databases and text files containing file directories,

metadata, and data documentation. Therefore, it is necessary to clarify these two key concepts that are uniquely important for digital archiving – metadata and data documentation – which are necessary for understanding how our digital archival content is organized, described, related, managed, and used within a Digital Archiving System.

What Is Metadata and Data Documentation?

Metadata is data – information about data, about the digital archival content. It is stored in a structured form suitable for software processing. Metadata is essentially equal to archival descriptions of digital content. Indeed, the descriptions of our content that we made in the previous stage will now, in the Digital Archiving System, become metadata, thereby adding to other types of metadata such as system-generated technical metadata or metadata on an item's access history. Metadata is therefore necessary for the goals of long-term preservation and access, as it allows us to maintain the integrity, quality, and usability of content.

Data documentation provides information about the *context* of our data, our digital archival content. It is often provided in a textual or other human-readable form. Data documentation in fact supplements metadata and provides information that enables others to use the archival content. For example, if we conduct a survey of victims and are preserving their filled-in questionnaires as our digital archival data, we should also preserve related data documentation (e.g., a document detailing the survey design and methodology). Given that data documentation is also "data about data," it could also be seen as a specific type of metadata, one which provides context and is recorded in human-friendly format.

Preparing Metadata and Data Documentation

While our digital files are safely stored and backed up on storage media awaiting ingest and archiving in the digital information system, we need to turn our attention to some housekeeping duties. They involve preparing our metadata and data documentation for the upcoming process to ensure the smooth ingest and proper archiving of files.

Image shared by CCJ, GIJTR partner organization in Colombia.

This involves having a clear and well-organized record of data documentation and metadata thus far in the process – what they contain and how they relate to one another. This includes tables/databases with lists (or directories) of file names, the files' metadata, and data documentation. Throughout previous chapters, we described how these documents are developed or generated through planning, inventory creation, review, selection, organization, description, and digitization of material. As a result, at this point in the process, we should have the following metadata and data documentation created:

A. **This document started its life as Identification Inventory and then, through processes of organization and description, grew into the Table of Archive's Structure. It contains metadata on the archive's structure, grouping of files in series, subseries, and folders, and additional descriptive and technical metadata we selected to put into it.**

B. **As a result of the digitization process, we have produced databases in which we recorded each digital surrogate we produced and the selected metadata about it.**

Further, digitizing equipment and software also generated additional databases with metadata we selected to capture technical attributes of the digital surrogates and/or history of actions on them throughout the digitization process.

Finally, we also might have produced text documents containing data documentation, information about the context of the digital surrogates we created, or the digitization process itself. This will allow others to understand how our data can be interpreted or used.

C. **A database of born-digital files for preservation with their basic metadata will either already exist or be easily created using simple software tools such as "DROID" or "IngestList."**

D. **There might be additional pre-existing tables/databases or text files containing metadata and/or data documentation about certain item groups or the entire collection.**

In order for our digital content, metadata, and data documentation to be properly ingested into the Digital Archiving System, we need to provide the system software with instructions on what these documents are and how they relate to each other. In this way, the system can, for example, correctly attach metadata in one database to the items metadata describes that are listed in a different database, and then to data documentation providing information about the given items' context.

As part of the preparations, we might also need to manually divide, merge, or combine some of our tables/databases to transform them into a more appropriate format.

The exact steps that we will need to take in this process in which we will need to prepare our metadata and data documentation, or how we will input information about their inter-relations into the Digital Archiving System, will depend on the characteristics of the archive and the system itself.

Yet, regardless of these specifics, we will always need to have a clear overview, a map, or a scheme of our metadata and data documentation and how they are related before we can begin with the ingest.

Preservation and Preparation of Data for Archiving

We can now move on to the preservation actions and preparation of our digital data for ingest and archiving.

Cleaning: The first thing we should always do before working with digital data intended for preservation is perform an antivirus scan by connecting the storage media to a previously scanned computer that is not connected to any local network or internet.

Backup: Then comes the backup. At the end of the digitization process, we already created backups of digital surrogates' master files. If we have not yet done the same for the born-digital data, we should create their backups now by producing two copies and storing them on separate storage media, if possible at two different locations.

File Naming: While our digital surrogates' files have already been named in line with the naming convention we developed and adopted, our born-digital files might still have their original names. We must therefore apply our naming convention to the born-digital files and name them accordingly. Their names will then contain the same components – identification, description, technical, or other – as those we selected and used for the digital surrogates in a way it was described in the digitization chapter. There are reasonably simple and easy-to-use software tools that can perform this task of renaming our digital files automatically within the parameters we set for it, such as "Rename Master" and "File Renamer Basic."

Metadata: In the previous section, we took stock of metadata and data documentation we collected thus far in the process. As explained there, we will need to ingest our metadata in a specific, fixed format that is recognizable by our Digital Archiving System. This specific format of metadata will be based on the metadata standard we selected to implement earlier in the process, and that we now need to apply for ingest of data into our Digital Archiving System.

If, as advised in this manual, in the planning phase, we have already made a decision on the standard we will apply for metadata collection and implemented it through description and digitization phases, then our metadata will already

have been gathered in accordance with that standard. Therefore, we should be able to arrange and prepare it for ingest in accordance with the system-recognizable format by making only basic technical arrangements or mapping our metadata to the standard. For example, in the digitization section, we mentioned that the so-called "Dublin Core" basic metadata standard is supported by most digital archiving software. Hence, if we applied this standard for the collection of metadata from the beginning, and we selected the software that supports it, we would now be able to translate the collected metadata into the format our Digital Archiving System can recognize and properly ingest.

Preservation of Metadata

In the earlier discussion of metadata and the importance of its proper collection and management, we mentioned the key role it has for long-term preservation of digital archival data.

This becomes even more salient and relevant at this point in the process, with the preparation for ingest and long-term preservation of our material. This is because, before we ingest and archive our data, we need to make sure that we capture the necessary metadata, which will allow our digital material to be adequately preserved, its authenticity maintained, and it remaining usable in the future. To understand which essential set of metadata we need to capture to preserve our invaluable data, we will need to get to know our digital files and their formats a bit better, including things such as our files' validity, quality, and fixity.

Identifying and Converting File Formats

Back in the digitization process, we established the need to store our digital material in file formats that are appropriate for long-term preservation. Primarily, these are formats that have a wide user/support community and are proven to be resilient to change over time. This is also why they are often called "lossless" as opposed to "lossy" formats that do tend to lose quality and/or change and degrade over time.

Our digitized material has already been stored in appropriate preservation formats through digitization, and now we need to make sure the same is true with our born-digital material.

We first need to identify the format of our born-digital files, which we can do with the assistance of specialized software, such as "DROID" or "Siegfried," that allows us to automatically identify the format of batches of our digital files. We will then proceed to change formats of the files for which we determine the need to be put into a different, preservation-appropriate format. Specialized software for conversion of files to different formats can be very useful in this process. Such software is format-specific (e.g., "Audio/Video to WAV Converter") which converts audio and video files to WAV format, or "CDS Convert," which allows conversion of documents, presentations, and images between different software formats.

The Importance of Using Proper Preservation Formats

Lossless formats, by rule, also produce larger files. Hence, for large collections and small organizations, such as CSOs, this can represent a challenge in terms of additional storage capacities they may require. However, this manual advises against making compromises with the selection of file formats, as use of proper preservation formats is essential for all following preservation actions and the success of the process as a whole

Validating Files

The next step in preparing our digital content for proper preservation in the Digital Archiving System is validation of our files – that is, establishing that they really are what we think they are.

In essence, through file validation, we check whether the format of a file is proper and correct – whether it is valid. Hence, through file format validation, we can check whether a file conforms to the file format specification – standards a specific file format such as .jpg, .doc., or TIFF must follow. As an illustration, file format validation

could be compared to the inspection of boxes or folders in a physical archive to ensure they are not damaged, otherwise items could fall out or be damaged.

In digital archiving, file format validation is particularly important for long-term preservation and access, for a number of reasons. Files with formats that are not valid are difficult to manage over time, especially when a file needs to be converted or migrated. Moreover, access might become difficult or impossible, as files with nonconforming formats become more difficult to open and use over time. Finally, files that are not valid will be more difficult – if not impossible – to render properly by future software.

Of course, we do not manually inspect whether a file format conforms to its specifications; there is software available to perform that function and identify and create reports on the files that are found not to be valid. We already mentioned one such software tool – JHOVE – in the chapter on quality control at the end of the digitization process, but there are also other tools, most of which are specialized for a certain group of formats.

Preservation Actions Should Immediately Follow Digitization

Preservation Actions Should Immediately Follow Digitization

File format validation and other preservation actions, along with the quality control procedures, should be performed immediately at the end of the digitization process either as an alternative or in addition to conducting them as part of the preparations for ingest, depending on a project's specific needs and workflow.

Fixity

Fixity, a crucial element of the long-term preservation of files as well as in maintaining their integrity, authenticity, and usability, means a state of being unchanged or permanent. In essence, fixity allows us to determine whether a file has been altered or corrupted over time and to track and record any such changes.

To be able to do this, we use fixity to record the initial state of a file before ingest by taking its "digital fingerprint." In fact, fixity software will record a number of a file's specific, technical characteristics and create an alphanumeric code – a "checksum." This checksum, just like fingerprints for humans, will be unique for that file and should not change over time. The checksum for a file will be recorded as part of its metadata so we can always perform the same fixity check and establish whether the file's checksum has changed – that is, whether a file has changed. Recording this type of preservation metadata is crucial for confirming and establishing a digital item's "chain of custody."

In addition to allowing us to establish any changes to a file that have occurred over time, fixity is also useful when we are migrating files between different storage media, units, or digital depositories. It is highly advisable to apply a fixity check after each such file transfer to establish any changes that might have occurred in the course of the file migration.

Further, fixity will allow us to verify that any copies of a file we create for backup are complete and correct. Fixity checksum can also be given to other potential file users so they are able to verify that they have received the correct file. There is a range of software that can perform fixity, such as "Checksum" and "Exact. File," just to name a few.

Quality Control

Many things can go wrong with digital files as they are created, managed, and stored before they reach the point of ingest. During digitization, due to an error or a virus, files can be damaged, made incomplete, or reduced in quality. It is therefore a good practice to perform as comprehensive a quality check of all our digital files as possible before their ingest and archiving. There is a whole set of tools that perform either specific or a group of quality control actions. Some examples include NARA's File Analyzer and Metadata Harvester, which has a range of functions, or, on the other side of the spectrum, a highly designed "Fingerdet," which helps detect fingerprints on digitized items.

Removing Duplicates and Weeding Files

While we are at it, we should use this opportunity to clean up our files a bit. Over the course of collecting, organizing, copying, and temporarily storing our digital files, it is likely that we will have created duplicates, or that folders contain hidden files or files that do not belong in them. Having duplicates and other unwanted files in our collection can create confusion, in addition to unnecessarily taking up space in our storage. It is therefore a good practice to remove them before ingest. Depending on the size of the collection, this could be a very time-consuming and error-prone task if performed manually. Luckily, there are software tools that can do this for us efficiently and reliably. Examples of dedicated tools for this purpose include “FolderMatch” and “CloneSpy.”

Metadata on Private, Sensitive, Confidential, or Copyrighted Data

Given the importance of data safety and security when archiving material related to human rights violations, it is highly advisable that, at this point, before the content is ingested, we make an additional review of the material with respect to privacy, sensitivity, confidentiality, and copyrights.

During the description processes, we should have already identified groups of materials or even single items that contain personal or sensitive information. Now we need to make sure all relevant metadata about such material is collected and appropriately linked to the items. Depending on the material and the archive’s access policy, it might be useful, or even necessary, to add further metadata here, specifically that which provides instructions for its future management regarding copyright, protection, or restricted access to the material.

Conveniently, there are standards and software that have been developed to provide assistance in that process.

Standards

Standards for metadata selection, collection, and use often include a full range of preservation metadata. Application of such metadata standards supports the preservation of digital items and ensures their long-term usability. A range of standards has been developed for handling preservation and metadata in general. As such a wide choice of options can often limit a clear view, we recommend an organization use as a starting point the "Preservation Metadata Implementation Strategies" (PREMIS) standard.

It should be noted that different metadata standards will often be integrated, or at least compatible, with software we use for metadata collection and management functions.

Resource Recommendation!

PREMIS has achieved the status of being the accepted international standard for preservation metadata. Both a strength and a limitation of the PREMIS standard is it must be tailored to meet the requirements of the specific context; it is not an off-the-shelf solution in the sense that an archive simply implements it directly to its data. Some of PREMIS's elements might not be relevant, and an organization may find that additional information beyond what is defined by the PREMIS standard is needed to support its requirements.

Software Tools

Thus far in this chapter we have mentioned examples of different software solutions that can perform specific preservation metadata collection and management functions, such as file identification, conversion, validity, and fixity checks. Such tools will indeed sometimes be designed to perform just one specific, or a group of similar, functions. However, these individual tools are also often used together as a more wide-ranging software solution, which can provide a full scope of preservation and metadata-related functions. Moreover, such multifunctional tools for metadata are then incorporated into comprehensive software solutions that can manage the entire process of digital archiving within a given Digital Archiving System.

In the planning section of this manual, where we discuss the selection of a software solution for our Digital Archiving System, we consider whether the option we choose has integrated support for the selected metadata standard, as well as all the necessary software tools to collect and manage preservation metadata to our archive's requirements. At that point, we could opt for an enterprise solution that provides an all-in-one option with all necessary standards and tools integrated into it. But an alternative would be to build a solution that meets our needs by using different, interoperable software, with each performing one of the preservation functions.

This stage of preparation of data for ingest and capturing preservation metadata makes salient the importance of our selection of the digital archiving software and the effect it has on the technologies and software tools we can and need to use. Therefore, the specific software tools we will apply in this phase, as well as later on, will fully depend on the type of solution we select for our digital archiving software.

Preparing the Digital Archiving System

Set-up and preparation of our digital archival system for its first ingest of digital files is a complex process that requires time, effort, patience, and reasonably advanced IT knowledge and skills.

Digital Archiving Systems cannot simply be installed and immediately used, as we do with standard commercial software. This is because any Digital Archiving System needs to be "instructed" on each and every aspect of its operations. Based on our requirements, we need to set the parameters in the system, create or design databases within it, create links between data and metadata, etc. Providing these "instructions" to our software might require anything from simply filling an electronic form or choosing an option from a drop-down menu to needing to use computer coding and other advanced IT skills.

The amount of time and expertise needed depends on the type of software solution selected for the Digital Archiving System. The rule of thumb we applied to the selection of software applies here as well. Commercial solutions will be simpler for both set-up and use, but will likely offer fewer options for adaptation. Open source solutions will mostly require more IT expertise and time – but can provide more suitable and tailored solutions.

Image shared by AVIPA, GIJTR partner organization in Guinea.

7.3 Ingest

This is the sweet spot, where the entire effort and process conducted so far comes together and results in the creation of our archive.

However, we should not imagine that we can just click a button, go have a tea, and return to see all our data, metadata, and data documentation ingested and properly connected to each other. Rather, the ingest process will need to be performed in parts by transferring material per group over a period of time. In the process, we will also likely encounter errors, discover incorrect specifications in a system, or similar that will need to be addressed, and the system will need to be fine-tuned and the ingest repeated.

After ingest of each group of material, we should produce at least one archival master copy of each item, at least two backup copies, and any derivative, working copies we might need.

Backup copies should be created and stored in line with the best practice rules described earlier (i.e., create multiple copies on two different storage media technologies and store them at different locations).

As a final step, we need to perform the same set of preservation actions we applied to our content in preparation for ingest. This includes scanning the material as well as all backup copies with antivirus software, and checking each file's fixity, validity, and quality assurance.

If we have covered the basics so far and made sure all the elements have been prepared, the process should be successful and we should now be able to enjoy the fruits of our work – our precious material previously scattered around the office and in storage units and basements – having been turned into a digital archive.

In the next step, we will take action to make sure the goals of our archive are also achieved – that it preserves our material for a long time and in a safe manner and provides as wide an access to its content as possible.

PART 8

6:00

ACCESS AND DATA SECURITY

Image shared by AVIPA, GIJTR partner organization in Guinea.

Now that we have finally created our digital archive, we could be forgiven for being eager to share it with its intended users and beneficiary communities, as we envisioned at the beginning of the process in our Guiding Principles.

However, providing access to any archive's content, and especially to a human rights violations archive, is not a simple or straightforward matter. Access is closely linked to preservation, but also to nearly all other processes and functions in a digital archive. Most closely, however, access is related to the data security function of an archive.

These two functions are also the focus of the third stage of a digital archive's life cycle: its safe opening to the world.

8.1 Access

Providing access to our archive's content is a balancing act between two of our Guiding Principles: defining the goal to provide as wide access to our archive as possible, and defining our responsibility to safeguard data and adhere to legal and ethical norms regarding data privacy, sensitivity, confidentiality, and copyrights. Having a well-considered and clear Access Plan will help achieve that balance.

An archive's Access Plan should guide both decision-making and implementation related to provision of access. There is no template for a digital archive Access Plan; however, we can identify five elements it should describe and define.

Access Objectives

While our archive's access-related goal might be generic – such as to provide wide access – the Access Plan should specify more concrete objectives that will contribute to that goal. For example, we could set an objective to create a set of useful finding aids to facilitate use. Or we could seek to make the access modes user-friendly and easily available.

Resource Recommendation!

Planning for Access

A systematic approach to planning different levels and modes of access is provided in detail in "Levels of Born-Digital Access" by the Digital Library Foundation (USA).

Specific objectives we set will of course differ for different archives, depending on their goals, users, content, etc. Regardless of these differences, setting up clear and concrete objectives will allow us to develop and then implement a comprehensive plan tailored to our needs and requirements.

Users and Modes of Use

Users of an archive have an essential importance for it – why it exists. The reason we aim to preserve our archival content long-term is to make it available to future users.

Therefore, our Access Plan needs to be grounded in users' needs and requirements. More than that, the Plan should envision a two-way relationship with users so that their input shapes the way the archive develops its access policies and practices.

On the primary level, we should differentiate between internal (archive and organization's staff) and external users. Within the internal user group, there will be varying levels of access, depending on a person's role and their access-related needs. Some staff members will have unrestricted access, while others might have restrictions in terms of different groups of material or type of access they have (e.g., to view or manage files).

When planning for external users' access, we can distinguish between fully open public access and access provided to predefined groups of users, such as registered users, members of selected external organizations, or similar. The Access Plan should define the level of access to be provided to each of these groups of external users.

Image shared by FAMDEGUA, GIJTR partner organization in Guatemala.

Regarding open, public access, the Plan should specify whether such access can be provided for selected groups of material in the digital archive, or for an entire collection. It should also define how material can be accessed (i.e., whether it is only available for viewing, or also for copying and reuse).

Access Levels

To provide tailored access to different groups of materials, we will need to first have them categorized based on their security status. We can do this using metadata collected in the description stage and in the phase of preparations for ingest. If collected properly, our metadata should allow us to clearly map any content that should be considered "sensitive." Archival data may be sensitive because of legal, security, or personal considerations. Marking material as "sensitive" may imply limited use (e.g., on-site only, or closed, or conditional access).

The number and names of access levels we create can vary depending on our needs but should cover the following three basic categories:

Open access: Open materials are available for use with no known restrictions. Materials can be directly accessed by users. Access may occur in an on-site public access point or online.

Conditional access: This refers to collections that include both open material and material that has restrictions. These restrictions may include materials that are deemed sensitive or under copyright. Conditional access is a continuum that includes documents with differing levels of restricted access.

Closed: Closed materials are not made available to users, but may eventually be made available after an embargo period. Collections or materials may be closed if they contain information protected by applicable law or private, privileged, or sensitive information.

Modes and Conditions of Access

Access to digital archival material can be provided in three main modes:

On-site access: Access to archive content is provided on a dedicated on-site public access computer with security measures implemented. This is a viewing-only mode of access, hence any form of copying of content is not allowed.

Controlled remote access: Secure remote access is provided to a limited group of users, either through a local computer network (LAN) or using a secure remote online access platform. This type of access is often provided to internal users who are not archival staff but work on archive-related projects. Also, this mode of access can be provided to partner organizations working on joint, archive-related projects or to selected individual researchers. The organization can specify whether this mode of access includes only viewing permissions or the users can also copy digital items.

Open access: Open access is provided on-site or through a website or dedicated online platform. Although open, access can still be controlled, for example by requiring future users to register before using the archive, submit a justified request for access approval, or similar. Open access allows for copying of archival digital material under the presumption that access is only provided to public material that allows reproduction.

Access Levels Scheme

Considering that different modes of access to material with varying permitted levels of access need to be provided for different user groups, to avoid confusion, it could be useful for an archive to create an Access Levels Scheme as part of the Access Plan. Such a scheme provides an overview of "who has access to what, and how" in the form of a table, such is the example shown in Figure 14. The scheme can be a useful tool for devising as well as for technical and logistical implementation of the planned access levels.

	Collection 1	Collection 2	Collection 3	Collection 4
Archive Staff	Open	Open	Open	Open
Project Staff	Open	Open	Conditional (view & copy)	Conditional (view only)
Partner Organizations Staff	Open	Open	Conditional (view & copy)	Closed
General Public	Open	Conditional	Conditional	Closed

Figure 14. Example of an Access Levels Scheme, with an overview of access levels for different groups of users and different groups of material

Opening the Open Level Access

It is useful here to make a distinction between a passive and active approach to provision of access. A passive access approach would be an archive created with the main goal of long-term preservation of the material for historical, legal, or other reasons. Provision of access might be of secondary concern for such an archive, and its efforts in this area might be limited to providing access only to requested materials or on-site only. The focus of such an archive would be to respond to users' requests and ensure it provides the appropriate level of access to the material for different user groups (e.g., institutions, researchers, etc.).

Most CSOs working with human rights violations archives, however, will likely be taking the other route of

Resource Recommendation!

Web Content Accessibility Guidelines (WCAG) **are an international standard that provides documentation and guidance on making online content more accessible to people with disabilities.**

an active approach to the provision of access, which is focused on facilitating and providing as wide an access as possible to its users.

The active access approach predominantly concerns "Open Access" mode and does not include materials marked with "Closed" access level. To make our "Open Access" mode truly open, we need to consider accessibility, searchability and usability of access to our archive and content. "Accessibility" concerns the ease of access to archives for everyone. For example, we should consider whether anyone with an internet connection – even an unstable or a weak one – can access the archive, how difficult it is to find and load the online access portal, whether it can be used via mobile devices, and similar. With respect to accessibility for persons with disabilities, we should consider providing a collection-level note about which born-digital materials comply with accessibility needs and/or what is required to render materials for those with visual or hearing impairments. Additional measures could introduce practical arrangements, such as a screen reader, color contrast, or adding tags to define reading order.

"Searchability" for our users determines how easy or difficult it is for them to find what they are looking for in our archive. Searchability of our archive will depend on the type and quality of metadata we collected about our material and how well we organize it and provide it to the archive's users. Using different metadata as "tags" or "keywords" associated with certain items or groups will help users find them more easily. Further, we can provide a user with a map of our archive to guide them by preparing a catalog using the descriptions of the collections, series, and other elements of our archive's structure. However, searchability of any online accessible archive will mainly rely on the quality of search that can be performed through a dedicated search engine.

"Usability" of an archive relates to how easy it is to use it. That includes, for example, how an archive's online access point looks and feels and what kind of user experience it creates. A well-designed and organized online access platform can attract more users and encourage current visitors to use it more. It can further support novel archive-related projects and extend the scope of both its use and its beneficiaries.

This is an especially important consideration for CSOs working with human rights violations archives, as their goal is often to not only provide access but also stimulate and facilitate different organizations and individuals to use archival material in their own projects, research, and activities.

Technology allows us to create a whole range of different online access platforms with a variety of formats, visual presentation forms, tools, and other useful features. These solutions can be impressive and made very attractive for users, which generates multiple benefits for both them and the archive.

However, we also need to bear in mind that any technical solution for an online access platform we might want to implement will need to be interoperable and compatible with our Digital Archiving System and any relevant external software tools we use.

Access Technologies and Tools

Providing a varied level of secure access to our content for different groups of users using different modes of access requires significant technological support that includes both hardware and software.

Thankfully – provided that we have, as suggested earlier in this manual, considered our future access provision needs at the time we were selecting our Digital Archiving System – we can now rely on it for the basic technology needed for the implementation of our Access Plan. For example, if we would have planned for the need to provide different levels of access to different users, both internally and externally, our Digital Archiving System would be able to provide support for it.

However, we will need to invest more time and resources in technical solutions, especially in terms of active approach to access, if we wish to build on these basic access capabilities. This would include using software and applications that allow for the development of digital archival tools and services users can benefit from, as well as improving the design, user-friendliness, and overall user experience of our online access platform.

The choice of software tools we will use for developing our online access platform will be fully dependent on our requirements – the type of platform we want to make, services it will provide, users it will target, etc. A useful tip in selecting software is to search online for an archival online access platform that looks similar to the one you want to develop and then work out which software and technologies were used to make it.

In addition to the technologies related to provision of Open Access, if we are planning to provide on-site or safe remote access, we will need to consider additional technological solutions. For on-site access, this would include a dedicated computer that is not connected to any local computer networks or the internet. We might also need other hardware or software for access to a specific group or format of the material. Safe remote access would also require specialized software that needs to be installed not only in the archive, administratively, but also by the users themselves on their devices.

8.2 Digital Archive Security

Protecting our invaluable collections and anyone who might be harmed by the misuse, altering, theft, or destruction of our archival content is an important topic for organizations working with archives documenting human rights violations. Any digital archive faces a wide scope of possible threats to integrity and protection of its content; the number and probability of threats are only increased for human rights archives. They range from threats to archival storage media through cyberattacks on an archive's information system and data to attempts to access data in an unauthorized way.

There are also legal and ethical obligations for all archives, which are only highlighted for those dealing with human rights violations material. They include protection of private, sensitive, confidential, and copyrighted data. For human rights archives, these are extremely serious obligations, not only because of the legal responsibilities they prescribe. Leakage or publication of a confidential or sensitive document, or unauthorized disclosure of a person's data, might bring related persons or organizations into dispute or even physical danger.

Digital Archive Security Plan

There are various aspects and elements to a digital archive's security, and to make sure we properly address them all, we need to make a Digital Archive Security Plan to guide us in devising security procedures and their implementation.

Image shared by FAMDEGUA, GIJTR partner organization in Guatemala.

While there is no universal template, a good way to approach security planning is to list and describe:

- Security-related obligations of the archive, based on the material it contains
- Security-related functions that the archive needs to perform
- Security-related actions that will be taken to ensure the functions are properly performed
- Tools and technologies needed for implementation

The descriptions of the security-related functions and obligations of the archive should be detailed and provide concrete information about the archive's requirements.

An additional segment of the Security Plan deals with different types of security levels for different material and groups of users. In essence, this mirrors the Access Plan and the Access Scheme, albeit from the security perspective; hence, we do not need to discuss this segment further. We will therefore focus on the planning of a digital archive's security-related responsibilities, functions, actions, and tools.

Security Responsibilities and Tasks

Our archive's main security-related functions arise from the security responsibilities we have with respect to our content, including:

Safeguarding data. The primary security-related responsibility of any archive is to ensure that its contents are not destroyed, changed, or stolen – that is, to safeguard its data. This includes protecting it from intended or unintended human-caused violations, as well as environmental damage, harm, or destruction.

Protecting personal data. This includes adhering to relevant national and international data privacy regulations relevant to a given archive. For human rights archives, this responsibility has an additional dimension, as violation of data privacy could have real and very negative consequences for the people or organizations affected.

Protecting confidential and sensitive data. With this obligation, again, human rights archives have an additional layer of ethical responsibility. Particular care and effort should be put into ensuring that any sensitive or confidential material is timely and properly identified and then also carefully protected in accordance with the developed procedures.

Protecting copyrighted data. The archive needs to follow and implement relevant regulations in the domain of copyright protection, relating to both access and use of its content.

Security Functions and Actions

System Protection is the first function of data security for a digital archive – its first line of defense. This is because, in order to protect the content (i.e., the data), we must first safeguard its repository. System protection includes safeguarding against system failures as well as protecting the Digital Archiving System from malicious acts of corruption or deletion. There is a wide range of information-security measures that can and should be taken to protect the Digital Archiving System from:

- Computer viruses
- Cyberattacks
- System failures and errors
- Inappropriate use or misuse of the system

The actions and measures to be taken, as well as the procedures that need to be developed to address each of these information security threats, will be highly specific for any given digital information system. What is universal is such planning needs to be conducted together with an organization's IT staff, with the assistance of external expertise (if necessary and possible). Security actions need to be well-designed and scheduled in advance. They should also include a plan for regular monitoring of any implemented information security measures.

Whether it relates to private, sensitive, confidential, or copyrighted data, data protection is the essence of our archive's security planning – the very reason we need it. There are three main instruments, or actions, an archive can take to protect its data.

A. **Access control and management:** Implementation of different levels of access for different user groups, potentially through different access modes, is the main action we can take to protect content that requires it. As explained previously, a precondition for developing and implementing successful access control is having quality metadata about our content, which allows us to identify material for which access needs to be controlled. In technical terms, this is implemented via a Digital Archiving System, through which we can specify different levels of access, monitor implementation, and record any violations of the rules.

B. **Redaction:** Redaction is the process of analyzing our archival content; identifying confidential, sensitive, or private information; and removing or replacing it. By redacting material in this way, we can make the non-redacted parts of an item openly accessible for our users.

   Frequently used redaction techniques include anonymization and pseudo-anonymization to remove personally identifiable information, as well as cleaning of authorship information. This is usually carried out by removing or replacing the sensitive/private/confidential information while retaining

the existing structure of the item in the version being provided to a user. Any redaction should always be made on a secondary copy of a file – never on the original, archival master file.

C. **Encryption:** Encryption is a computer technique that protects digital material by converting it into an incomprehensible, scrambled form. An encryption key is then created that needs to be used to unscramble the data and convert it back to original. Encryption can be applied at different levels, from a single file to an entire hard disk.

However, encryption also adds to the complexity of the digital archiving process and should therefore be avoided if possible for archival copies. It is only effective when a third party does not have access to the encryption key in use, which is why the key needs to be safely stored and protected. The loss or destruction of these keys would result in data becoming inaccessible. Encryption must also be actively managed and updated to remain secure, since it can lose its effectiveness over time.

Security Technologies and Tools

In protecting and securing the Digital Archiving System and its software and applications, we will need to apply a number of solutions related to different areas of information security. These solutions need to be designed by experts in this area. Regarding the technologies involved in our archival data protection, different software tools can be helpful for each security action.

Resource Recommendation!

With respect to redaction and encryption, there are several software options available, some for specific contexts and functions. Useful lists of such software resources can be found here (for redaction) and here (for encryption).

Documents before organizing the archive, image shared by FAMDEGUA, GIJTR partner organization in Guatemala

PART 9

9:00

MAINTENANCE: PRESERVATION, DEVELOPMENT, AND MIGRATIONS

We have reached the final stage of the first cycle of our digital archive's life. We have planned out and created our digital archive, preserved and secured our invaluable archival content, and made it accessible to the community.

However, we cannot just sit back now and do nothing. The wheels of a digital archive need a lot of continuous oiling to keep preserving and making our content accessible. This is why in digital archiving in particular, the notion and approach of active maintenance is essential; it is an integral part of long-term preservation.

Here is how the United Nations views the matter:

Digital Preservation is the active management and maintenance of digital objects so they can be accessed and used by future users.

The goal of digital preservation is the accurate rendering of authenticated content over time to ensure its authenticity, accessibility and usability.

Active maintenance is not only needed to keep our digital archive running, but also the key to long-term preservation, authenticity, and access to digital content. If the format of our files becomes obsolete, our storage media fails, or our backup software is outdated and flawed, our invaluable digital content may become compromised, damaged, or lost altogether, along with all the work we have put into building and developing the digital archive.

Given that active maintenance is not an afterthought but rather the very core of digital archiving, it requires a systematic approach and regular performance of a set of actions that include monitoring and migration. As with access and security, it is a good practice to create a Maintenance Plan centered around the two main sets of maintenance actions: listing, describing, and scheduling the execution of maintenance **monitoring** and **migration** activities. The specific elements of the Maintenance Plan, such as time periods for regular checkups, or concrete procedures, will be dictated by the circumstances of a given archive. We can, however, describe the key elements and actions that need to be included.

9.1 Active Maintenance: Monitoring

To properly maintain our digital archive, we need to monitor its functions and elements and make the necessary adjustments. This primarily includes monitoring, checkup, and preservation actions on data, software, and storage media. In addition, we need to regularly observe, revise, and update our data security and access plans and their implementation.

Data Monitoring and Preservation Actions

Monitoring and preservation actions we need to conduct on our data in the maintenance phase are in essence a continuation of the work we performed as part of the preparation of material for ingest – from plain backup of data to checkups of file format, validity, fixity, and quality assurance.

At this stage, we need to plan and schedule regular, periodic performance of these preservation actions, check for any irregularities, and then follow up to amend them. We also need to plan for these functions as a necessary step in any major archival data-related activity, such as data migration or software replacement.

Antivirus checkups. As always, ensuring that data is virus-free is an essential precondition for any further actions on data. In addition to antivirus measures in place for the entire Digital Archiving System, we also need to be mindful and perform antivirus checks on our digital content whenever it has been exposed to a networked environment or other virus-related threats.

Backup. For the maintenance phase, it is important that our backup copies are also monitored and replaced when appropriate – mirroring any actions on our archival master files. Hence, backup files should be subjected to the same type of scheduled checkups as are our master files. Alternatively, we could make new backup copies from master files following their regular checkups. Additionally, if there are any changes to the archival master files, their backup files will also need to be replaced.

It is a good practice to back up considerations in the maintenance phase whenever possible, including development of a so-called "Disaster Recovery Plan." This refers to creating a plan on how our data will be recovered or replaced in case of any major natural or human-caused

Image shared by FAMDEGUA, GIJTR partner organization in Guatemala.

failure, damage, theft, or malicious attack on our digital archival content or system. The plan will be based on our existing backup arrangements, which define the number of backup copies, their geographic location, and type of storage media used, as described earlier. The disaster recovery plan should provide instructions on which of our backup copies should be used, in which disaster type circumstances, and by which technological means to replace and recover our data.

Format and obsolesce. In the pre-ingest phase, we made sure that all our files are in formats that are operational and can be opened and properly displayed by currently available software. Similarly, in the maintenance phase, we need to schedule regular audits of our file formats to ensure that they are not in danger of becoming obsolete. If we find that a format we are using is becoming obsolete, or that support for it will be discontinued, we need to act, which usually means migrating files to a newer or more suitable format. For both file format monitoring and migration we can use specialized software tools, some of which we already mentioned.

Fixity, Validity and Quality Assurance. Even preserved digital files can change over time, which can then affect their format and/or quality. Hence, similar to monitoring file formats, we need to also plan and schedule regular checkups of our files' fixity, validity, and quality. Equally, we

should plan to include these preservation actions as part of any major data-related actions, such as data migration.

What differs is how we will approach irregularities we might detect or any changes we find to our data. In case of a detected change in a file's format, quality, or fixity, we can follow a three-step rule of thumb: repair, restore, and document.

This means that in the first instance we can try to **repair** the file using dedicated software tools for the given file format. If repair is not an option, we should **restore** the file from one of our backups. In case we do not have a backup or it is not usable, we can decide to preserve the

Image shared by FAMDEGUA, GIJTR partner organization in Guatemala.

changed original file. Regardless of what we decide, in the end, we need to **document** our action and detail the decision that was made and why, to be preserved as part of metadata along with the file.

Monitoring Software

The software we apply in our Digital Archiving System – be it open source or commercial, an all-in-one solution or a combination of specialized tools – also needs to be regularly monitored so that it continues to meet our requirements and avoids becoming obsolete.

An archive's requirements, as mentioned earlier, are not "set in stone." As it is expected that they will change over time, we need our software to support those changes – which is why it is so important for the software we use to have a strong support. We can then rely on this support – in the form of a community of software users and developers or a commercial service – to provide upgrades or additions for any new or revised requirements.

Monitoring and improving our software will ensure that it continues to meet our requirements, even when those requirements change. However, if our monitoring shows that a specific software can no longer be adapted, or that it is losing its support community, we need to find a new appropriate software solution and migrate to it.

Resource Recommendation!

Software Tool Registers

Although digital archiving would not be possible without software tools, and their quick pace of proliferation has been very beneficial, the sheer number and scope of possible and offered solutions can create difficulties finding and selecting the right option. A number of digital archiving software registries have been created, which provide lists and descriptions of different tools. A good starting point is the COPTR registry, which has the advantage of drawing information from a variety of sources and thus provides a good overview.

Therefore, an important element of monitoring our software as part of maintenance is to follow the new developments and services provided through upgrades and novel solutions, and to have access to a community of software users and developers. This is especially the case for civil society-run human rights archives, as many of them struggle with expertise, resources, and capacity needed for development and maintenance of the technological element of a digital archive. There are inspiring examples that show how such synergistic partnerships can be built, and new out-of-the-box solutions can be applied to shared technology-related problems.

Monitoring Storage Media. Monitoring our storage media is necessary to detect any errors or damage in a timely manner, as well as to prevent it becoming obsolete or outdated.

Over time, storage media can become unstable and unreliable, and cause data corruption or loss. A rule of thumb for a safe preservation practice is for storage media to be given only a short lifetime, sometimes estimated at only three to five years. This means that after this period we will need to find and obtain a new storage media and migrate our data to it. This migration is somewhat less demanding than the file format migration but still requires all data preservation actions to be performed as part of the process.

It is a good practice to expect and plan for failures – both human – and technology-caused – to happen to our storage media over time, even in the best of circumstances, and regardless how good the technology is. This is why developing a clear Disaster Recovery Plan is so beneficial for a digital archive's maintenance.

The best strategy, however, remains to create a strong, resilient backup system with multiple independent copies, stored in different locations and using different technologies (whenever possible). Coupled with regular performance of all data preservation actions, risks will be minimized and spread, thus ensuring we never have to rely on a single piece of technology to preserve our invaluable data.

Monitoring Access and Data Security

Implementation of our Access and Data Security plans also needs to be regularly audited to ensure they remain functional and meet the requirements. When requirements change or shortcomings are identified, the plans and related practices should be revised.

In terms of access, monitoring includes following the statistics of our data use and users. Such data should be provided by the access software solution we use, and can help us better tailor, organize, and deliver our access services.

With regard to security, we must rigorously monitor the planned and implemented arrangements to ensure timely identification of any weak spots or shortcomings that might put our data under threat. Data on access and the use of the archival material can also be useful for reviewing and revising our Data Security Plan.

We should also always be on the lookout – when possible, through a community of users or other CSOs – to advance our access services and security arrangements by applying novel technologies that can sometimes substantially improve both the user experience and the safety of our data.

9.2 Active Maintenance: Migration

Migration is one of the concepts that are uniquely important for digital archiving. In a sense, it represents the very essence of it: the constant change, adaptation, and solution finding required to keep our digital content alive, preserve it, and make it accessible in the future.

Ominously, this comes at the very end of a digital archives' life cycle, signaling the end of one of its iterations and the beginning of another.

Migration of Data, Software, and Storage Media

In the previous section, we introduced different ways in which migration is an essential activity in the maintenance of a digital archive, as it allows us to preserve our data regardless of its format, the software we use, or the storage media.

Making a decision to migrate data, software, or storage media should be planned and scheduled to the greatest extent possible. It should also not be made easily. New, groundbreaking software or hardware solutions might be tempting as a great way to improve our archive's services, but we should be wary of untested solutions – and bear in mind that any migration is not a simple process, as it requires time and resources, as well as inevitably brings its share

of risks. Even in a simple transfer, and more so in a complex format or software migration, data can be changed, damaged, or lost. However, if we hesitate for too long, the software might become obsolete and make the task of migration much more difficult.

Therefore, migration needs to be performed timely, in a systematic and carefully planned way, following certain basic rules and best practices.

- Whether we are migrating data, software, or storage media, we should always include a set of preservation actions – including checkups of fixity, validity, and quality assurance – as a mandatory step both before and after the actual transfer of files.
- For file format migration, it is important to always retain the original file along with the new migrated file, as the migrated file might have lost some of the properties of the original file, which is not always detectable by the software. In such a case, we will have to decide what data we consider acceptable to lose (if any).
- To reduce the risk of changes to files during file format migration going unnoticed by our software, as part of a migration process, it would be a good idea to plan for a quality control test. This would include manually opening and checking a reasonably sized sample of migrated files based on a set of acceptance criteria we develop – for example, in terms of its formatting, look and feel, and functionality.
- For software and storage media migration, it is also a good practice to retain the original files for an appropriate period of time following the migration – anything from a few days up to a year or longer, depending on the type of migration. This is because, frequently, it is only post hoc that we discover process shortcomings or data changes that occurred during the migration. If we are then still retaining the originals, we can simply repeat the process and avoid data change.
- We should always include backup copies in any migration plans and workflows, and make sure that after the migration is completed, new backup copies are created from the migrated master files.

PART 10

0:00

A new collection of material has been gathered for archiving and long-term preservation, raising a set of new requirements. We will need to migrate our content to a new Digital Archiving System. In other words, it is time to start the process anew.

MEANS AND RESOURCES FOR BUILDING A DIGITAL ARCHIVE

If reading the manual so far has made you think, "This is all very nice, but my organization does not have anything close to the resources, expertise, or capacity to handle this entire exercise," you are not alone. In fact, you are likely in a significant majority.

The good news is this is no good reason to close this book and give up!

If the content you are working with has historical and social value, if it can facilitate transitional justice processes and potentially support positive social change, and if your organization is dedicated – and has a strong organizational will – to do its utmost to preserve that material, then you should not walk away from that cause just because the degree of difficulty is high. The following section describes why and how it is possible to build and maintain a digital archive that's scaled to your organization's capacities, even when your initial resources and expertise might be modest.

10.1 Resourcing

The main resource the organization will need to have in excessive amounts throughout the process of building a digital archive is a strong commitment, along with persistence and willingness to address a variety of challenges.

Such a strong and comprehensive commitment is the most valuable asset in the digital archiving process, and the closest thing to a guarantee of eventual success. This is because, even under the best of circumstances, building a digital archive is a process of continuous problem-solving of constant challenges that demand new and creative solutions. Even if at the beginning an organization has all the resources needed to build a digital archive, if it lacks commitment, it may as well give up, as it will be faced with continuous challenges and demands for maintenance, migration, and transformation, which never end in digital archiving.

"When there is a will, there is a way," on the other hand, is the right way to approach digital archiving. An organization can start small and plan modestly

Image shared by CCJ, GIJTR partner organization in Colombia.

without overextending itself, giving itself the time and space to develop the digital archive. And even when digitization efforts are seemingly minor, they are likely to prove useful. Such efforts accumulate and incrementally build different functions and qualities of an archive.

"Doing something is way better than doing nothing" is therefore the second motto that should be applied by CSOs looking to start their digital archiving efforts. While this manual describes an optimal process of building and developing a digital archive using a sequence of steps, in reality, companies do not operate in a vacuum, so doing so may not be workable for a given firm.

An organization might need to urgently preserve and safely back up a portion of its digital material. Or maybe it has boxes of important documents that are partly damaged and need to be digitized and preserved immediately. Addressing such urgent needs would then likely involve a small and localized digitization or digital-preservation effort, perhaps using a couple hard disks and an Excel spreadsheet. Yet, even such small efforts often prove very useful and meaningful

later on – and can also create positive momentum for further development of the digital archiving function of the organization.

"It is a battle for the hearts and minds." Building a digital archive requires the right people – and many of them – to be on board and to support the mission and goals of the archive. This includes both staff and external actors. Internally, it is not enough for directors or managers to agree. Virtually all staff from administration to IT need to support the decision to build a digital archive and perform the necessary implementation in the long term. Equally, an organization's key external partners, friendly CSOs, experts, journalists, donors, and mentors should all be supportive – and, to the greatest extent possible, be engaged in the process of building the archive. Having a strong network of such relevant actors supporting the organization's digital archiving efforts is an invaluable resource in this process.

Resourcing and Fundraising Plan

Throughout this manual, we identify and point out a wide range of financial, technical, and organizational resources that are required for building a digital archive. This includes human resources and expertise, financial resources, computers and servers, storage media and cloud services, digital archiving software, other specialized software and hardware, various scanning equipment, office and archive space, internet access, furniture, and logistics.

This list above can seem overwhelming – and it is likely not even comprehensive. However, again, this is not a reason to panic and walk away. These requirements relate to different phases of the digital archive's life cycle, hence the resources can and should be obtained and built gradually. As already noted, but worth repeating, an organization can set its own pace and plan the time frame for building the archive in accordance with its capacities and ability to obtain the necessary resources.

That process starts with planning. The organization should conduct a thorough assessment in consultation with experts (if possible) of the overall resource needs related to developing its digital archive, broken down by phases and activities. This should include an analysis of the more immediate, short- and medium-term needs in developing the digital archive (i.e., which resources need to be obtained

and within which time frame). Additional assessments of the resources the organization already has, or can easily acquire, should then be made.

Based on these assessments, a realistic, feasible plan should be made for securing the necessary resources and funds for developing a digital archive, both in terms of immediate steps and for the long term. Such a Resource and Fundraising Plan should be developed as an open and flexible document, with a view of updating and improving it as new opportunities, potential partnerships, or priority needs arise.

Fundraising is far from being the only game in town when it comes to digital archiving resourcing. This is because one-off financial donations cannot resolve resourcing issues for a digital archive for the long term. Rather, these donations require a combination of different sources of support and funding, and a "pallet" of varied approaches to acquiring that support and resources.

For example, large IT companies such as Google have programs through which they support nonprofit organizations worldwide. This can include in-kind donations of anything from hardware and software to services, training, and/or expertise. Further, through fruitful partnerships with other CSOs, or indeed other institutions such as universities or media, human rights CSOs can far extend the scope and number of potential donors or supporters they can reach.

There are also ways to obtain some of the necessary resources or equipment for digital archiving that do not require financial means. Open source software is certainly the main such element in a digital archive resourcing puzzle, as it can provide a service that's superior to commercial solutions – free of financial cost, although the investment in time and human resources will then need to be substantial.

Setting up and running an internship or volunteering program attached to the archive could be a way to add strength to the archive's human resources – which can be used beyond simply the repetitive work needed to perform the many time-consuming archival and technical tasks, such as description or digitization. Interns and volunteers bring their own and valuable knowledge, experience, ideas, and even expertise, representing a substantial addition to a digital archive's resource portfolio.

Donors as Partners

In building a digital archive, we are looking for both long-term donors and partners. In fact, for the most part, donors need to be partners in digital archiving, as it is necessary that they fully understand both the importance and demands of the process. Such relationships also need to be built on mutual trust in order to keep them going long term. Donors should always be informed of any major difficulties or problems we encounter, as well as the successes.

Therefore, in approaching the donors – both existing and new – for archive funding, we should be open about the need for continued support, as well as realizing that the process is complex and will take time.

However, we should not seek their support for building or maintaining the digital archive, but rather for what we plan and want to do with the archive and its content. We should approach the donors with a proposal to build a virtual museum for victims, not with a request for the new software we need for that museum. We should propose to potential supporters the creation of a searchable database of disappeared persons, not request funding to hire another programmer.

Finally, as mentioned in the section on outreach, visibility and presence in the community help. Making the archive's benefits salient and clear – to not only the donors and partners but also, whenever possible, a wider public – adds a valuable layer of social relevance and credibility to the archive. Building awareness of the value our digital archive has for the wider community is an important – if not the most important – asset for obtaining resources and funding, especially in the long term.

A highly recommended approach, therefore, is to make use of the archive for visibility purposes as early in the process as possible. This can include creating public-facing and outreach products using archival material and/or providing access to at least a segment of the future archival content. This would give us a palpable product that can provide a taste of what our archive can deliver, and make it much easier to attract and retain donors' attention.

10.2 Capacity Building and Networking

Because you are reading this manual, you are already engaging in the two activities from the title of this heading: building capacity. Indeed, this publication, as well as the project it is a part of, have been designed precisely to address the need for capacity building and connecting CSOs working on the digital archiving of human rights violations. This includes CSOs sharing digital archiving lessons through interviews and questionnaires, discussions held during a virtual conference in 2022, relevant literature, and a collected set of informative resources made available as part of this project and available on the GIJTR website.

The importance of and benefits from connecting with like-minded partners, as well as from the gradual development of capacities of the organization in digital archiving, are not always tangible. Their value and importance is not exclusive to raising funds and obtaining the necessary resources for digital archiving – although those are certainly important as well. Networking and capacity building are also crucial for the long-term development of the archive and the organization, as well as for successful preservation and effective utilization of our digital content.

In terms of capacity building, this relates primarily to extending and specializing an organization's human resources, as well as strengthening the organizational aspects that support digital archiving. The long-term goal is to build expertise within the organization. This can be achieved through staff training and new hirings, by adding workforce when needed, by specializing staff members for certain tasks and roles related to building digital archives, and through education of all relevant staff on key archival principles and practices. It is also important to be mindful that this process of building organizational capacities for digital archiving needs to be backed with adequate administrative and logistical support from all parts of the organization in terms of management of finances, human resources, establishing new procedures, procurement, etc.

Considering that resources for training and staff development in CSOs' budgets are, to put it mildly, rarely substantial, it is highly important to note that opportunities exist for free or sponsored training and education on digital archiving. Literature is also prolific and widely available, as well as digital archiving software users and developers and access to communities of digital

archivists. There are also programs that provide for sponsored, external expert assistance for CSOs in the process of building a digital archive. Such experts can not only guide the organization's staff in developing the archive, but also provide training and later assistance as long-term consultants with in-depth knowledge of the archive.

Finally, an organization does not need to do everything by itself. Sometimes, the demands of a task or function will simply be overwhelming and out of reach for an organization, even in the longer term. Or an organization can make a considered decision not to invest its resources in a certain complex or resource-draining aspect of the digital archiving process. This could, for example, be the case with digitization, the development of an access platform, or implementation of data security software. In such a case, we can decide to entrust another organization with performing this segment of work.

Outsourcing parts or even the majority of digital archiving work can be an excellent option, especially for small organizations and those working in conflict-affected areas or under authoritarian regimes. The goal is then to find a trusted partner organization that is willing and able to perform this element of work. Often, smaller CSOs will work with the larger ones that have more capacity and can provide outsourcing for certain services.

Similarly, a number of CSOs will join in a partnership or network, thereby merging and multiplying their capacities. Universities can also be a good choice for a partner, as they are trusted institutions that can be expected to provide long-term support.

Networking, and especially joining forces with external partners, could also be seen as a way to build the necessary capacities of an organization to build and preserve a digital archive. The benefits are numerous and include finding solutions for shared challenges, and merging forces to either obtain a certain resource or jointly address a particular issue.

It is a good practice for an organization to keep track and make use of different networking opportunities, such as virtual and regular conferences; online digital archiving communities; public promotions of the organization's work in relation to the archive, which then create new connections; or close monitoring of the field, along with keeping an eye open for potential synergies.

PART 11

OUTREACH AND SOCIAL ACTIVISM: ARCHIVE IN ACTION

This is it.

This is the moment you have been working for all along in the creation of the digital archive.

Of course, you want to preserve the content and save it for future generations. That is the ultimate goal. But, in terms of the people here and now, people who might have inspired you to create the digital archive in the first place – victims seeking ways to document their suffering and realize their right to reparations, or children wanting to know more about their disappeared parents, or a prosecutor who needs a credible source of evidence to start a war crimes case – this is it.

Outreach Strategy

Now we can start developing those portals and online museums, planning our educational workshops and publications, inviting researchers and lawyers to start testing our search engines, for which we worked so hard to collect the metadata.

The ideas can overflow, especially once the archive has been built and the enthusiasm is high. It is important, however, to be aware that the organizational resources are limited and require careful planning for their most efficient use and the highest impact of the archive's outreach efforts.

Therefore, developing an Outreach Strategy that describes your outreach goals, beneficiaries, and plan of activities – one that can pay back multiple dividends – would clearly be a good investment of time. Outreach and other external activities of the archive are not only a way to achieve our activism goals – be it to inform, educate, raise awareness or seek truth, accountability, or reparations – but also to make them more feasible and attainable. An archive with a bigger presence and impact in the community, with wider visibility and credibility, will be more likely to engage a whole range of actors necessary for its long-term sustainability. This includes expert staff, users, CSOs, and other partners, donors, teachers, students, universities, supporters, mentors, journalists, and others who can and need to contribute to a digital archive's successful creation, maintenance, and development; to its continuing life cycle.

Image shared by FAMDEGUA, GIJTR partner organization in Guatemala.

Outreach Goals, Actions, and Beneficiaries

Any archive's Outreach Strategy will clearly be highly specific, reflecting its unique context, content, and priorities. However, there are joint elements common to the human rights archives' approach to outreach goals, target groups, and actions.

Goals

It is important to set our priorities, and to do so clearly. We might want to do *everything* with our archive, but that might lead us to achieving *nothing*. Depending on the organization's capacity and resources, we might have more than one goal, but in order to keep our outreach work focused, we should not have too many.

When there is a clear priority, a major event, or a process in the community that would greatly benefit from our archive's involvement, we should make it a clear priority and focus our resources on it. For example, if a country is in a post-conflict period and currently establishing a state sponsored truth-commission, the primary outreach concern should be clearly focused on the immediate effort to reach out to victims and the commission, and eventually provide support and archival material for the truth commission's work.

The strategy should also make clear how our goals will be achieved – to operationalize them – to determine which concrete objectives need to be met to realize the goal(s).

To continue our example, our goal in this case might be to provide credible support to the work of the truth commission through provision of archival material. This goal could be achieved by meeting a number of objectives (e.g., reaching out and establishing partnership/cooperation with the commission and/or those appearing before it; developing organizational capacity to quickly search for, locate, and provide access to/copy of relevant material; and similar).

The objectives would then guide the development and tailoring of relevant outreach activities.

In this case, these could be educational workshops on the use of archives for victims and truth commission representatives, or hiring additional staff and developing a workflow for the quick provision of archival material to the commission.

Activities

If any general recommendations can be given with respect to development of outreach and socially engaged activities, it is to not be general, and not be led by templates. On the contrary, a human rights digital archive's Outreach Strategy needs to be localized and tailor-made to listen and respond to the community to create ready-to-use, out-of-the-box solutions and channel its activities and messages through modes and media that can best communicate them. By doing so, it will achieve the most impact in the immediate community in which it operates.

Beneficiaries

In a different archive's Outreach Strategy, this heading might read "Target Group" or "Audience." However, those terms imply a passive status of a strategy's beneficiaries, a one-directional communication from the archive to a specific group an organization wants to reach and to which it seeks to deliver its message or information.

Such an approach, however, would not be suitable for an Outreach Strategy for a human rights archive. Beneficiaries not only are the recipients of our messages, material, publications, workshops, and similar outreach products or services we offer, but also benefit from the archive's outreach efforts – not through mere reception, but through being empowered and taking an active approach, engaging with and contributing to the achievement of the archive's goals.

An Outreach Strategy for a human rights archive should therefore stimulate and support beneficiaries to use the archive to create their own projects. This should include a broad range of social actors – from victims and their families to artists and academics – who can all use their voices to tell the stories our archives contain.

There will always be more stories to tell from our archive that are worth it for a society to hear. The more and different actors and beneficiaries of our archive we can attract and engage, the louder and further their voices will be heard – and the more value and impact our digital archives will eventually have.

PART 12

SUMMARY

The purpose of this manual is to provide an informative and practical guide to digital archiving for CSOs working in transitional justice and human rights, both those beginning this process as well as CSOs working to maintain and develop their digital archives.

To address the growing need of CSOs for information resources, capacity building, and networking in the field of digital archiving, in 2021, five organizations – members of the GIJTR consortium – initiated the project "Supporting CSOs in Digital Archiving." This manual is a result of this project. It was conceptualized, devised, and developed on the basis of research, exchanges, and knowledge gathered throughout this two-year project. The manual is grounded in, and built on, the experiences and insights of more than 40 CSOs from 24 countries that provided their input and participated in this project.

For CSOs that have gathered or obtained invaluable collections of materials that are – in archival terms – undefined, unorganized, and therefore unusable, transforming these collections into proper, usable, and sustainable archival collections becomes a necessity. When these archives contain physical material that needs to be digitized, born-digital, or both, achieving this goal will require an organization to embark on the process of digital archiving.

This manual provides a guide for the development of a digital archive, starting from the decision to create a Digital Archiving System through its establishment and development to its continuous change, adaptation, and maintenance. It applies the Digital Archiving Life Cycle Model, which reflects the key characteristic of digital archiving: its circular and continual character, a cyclical process in which stages follow one after the other continuously, without an end point, like hours on the clock. A new cycle begins at the end of the previous one, starting a new iteration of a digital archive development. The Life Cycle Model also draws attention to the need for taking action and actively managing a digital archive throughout its life cycle and presents the wide scope of responsibilities involved in the digital archiving process.

0:00

Planning and Organizing Stage

As a first step, we need to develop a General Plan, which will define the Guiding Principles of the archive, as well as address key organizational, technological, and resources-related issues that will be encountered throughout the digital archive's life cycle. The Guiding Principles are defined based on responses the organization gives to a set of core questions, such as, What needs to be preserved? Why? Who will use it, and how?

The General Plan needs to be complemented with the creation of an Identification Inventory, selection, organization, and description of the material we want to preserve. This is because any further decision or action in the process will rely on information about the format, amount, scope, size, topic, or other characteristics of the collected material for preservation, as well as its ability to identify, manage, and locate groups or individual items.

To round off this stage, we will need to plan, design, and select our future Digital Archival System – a digital repository and content management system that will be hosting our archival content. A Digital Archiving System consists of hardware and software elements, which we will need to carefully select at this point, given that their characteristics will affect other important aspects of our digital archive.

3:00

Digitization and Preservation Stage

The second stage includes a group of digitization, description, preparation, and preservation actions, which all lead to the process of ingesting our digital material into a Digital Archiving System. These actions are separate but go hand in hand, as they are interrelated and

need to be well-coordinated. Digitization of any physical material needs to be done in sync with the decisions on how these objects will be described (i.e., which information, or metadata, about them needs to be captured in the digitization process). It is similar with born-digital material, as its metadata needs to be selected as well.

This is followed by a number of actions aimed at proper preservation of the archive's content by maintaining its integrity and credibility (i.e., ensuring that the objects are not compromised and any changes made to them are recorded). The material, both digitized and born-digital, is then ingested into the Digital Archiving System and onto the storage media. In this process, the content and its descriptions – its metadata – are captured and stored in the Digital Archiving System. Additional checkups are then performed and backup copies are created and stored separately.

6:00

Access and Security Stage

Providing access and data security is the main task in the third stage of digital archiving. These two separate functions are interrelated and need to be kept in balance to provide for the optimal effect – the widest possible access to be provided – while maintaining data safety and protection. This includes protection of any private, sensitive, or copyrighted data, but also measures to provide for safety of data and storage systems against physical harm and cyber-threats.

Providing wider access (e.g., by making a digital archive accessible through open databases or online platforms) will pose an additional set of data security issues compared with providing access to a closed group of users. Similarly, different items in the digital archive may require varying levels of protection and controlled access. Therefore, appropriate levels of access need to be defined for different groups of users in relation to different parts of the archive.

9:00

Maintenance Stage

Maintenance through preservation and migrations – these are the actions that dominate the fourth stage of digital archiving. Once the digital archive has been designed, set up, and populated – its data preserved, secured, and made accessible – all these functions need to be maintained and monitored and the content and the system need to be managed and eventually migrated and transformed. Regular maintenance checks need to be performed on the data (to ensure its continued integrity and credibility as well as format usability), system (to provide for continued security and open access), and hardware and software technologies (to ensure their proper functioning and act timely when they need to be migrated or transformed to prevent them from becoming obsolete).

0:00

A new iteration of the digital archiving process begins.

The manual argues and describes why and how it is possible to build and maintain a digital archive scaled to any organizations' capacities, even when initial resources and knowledge might be modest. The main initial and main resource the organization will need to have in excessive amounts throughout the process of building a digital archive is a strong commitment, along with persistence and a willingness to address a variety of challenges.

Funds can be raised, capacities can be built, networks of partners and donors can be created – and there is a pallet of possible resources a CSO can draw on in developing their digital archive. Developing a Resourcing and Fundraising Plan allows for the creation of a realistic,

time- and resources-wise feasible framework for securing the necessary funds, capacities, and technologies for developing a digital archive, both in terms of immediate steps and in the long term.

Developing an Outreach Strategy, including goals, beneficiaries, and a plan of activities, and conducting extensive outreach efforts is not only a way to achieve our activism goals – be it to inform, educate, raise awareness, or seek truth, accountability or reparations – but also a way to make them more feasible and attainable. An archive with a bigger presence and impact in the community, with wider visibility and credibility, will be more likely to engage a whole range of actors necessary for its long-term sustainability. This includes experts, users, CSOs, and other partners, donors, teachers, students, universities, supporters, mentors, journalists, and others who can and need to contribute to a digital archive's successful creation, maintenance, and development – to its continuing life cycle.

The difficulties along the process of creating and developing a digital archive are many and require patience, perseverance, and the readiness of the organization to invest substantial resources and effort – and progress should be measured in small steps. Yet, given that the goal of preserving the painstakingly collected material and giving life to the memory it records is so important and valuable that the trade-off is clear, the seriously challenging process of digital archiving is well worth taking for any CSO with an invaluable archive to preserve.

PART 13

GLOSSARY OF KEY TERMS AND CONCEPTS[2]

Digital Archiving is the process of creating, managing, and developing digital archives.

Digital Archives are archives that contain material in a digital form – including both born-digital items and digitized versions of originally physical material – stored on digital media and managed through digital tools. Just as in physical archives, the material in digital archives is organized, described, arranged, stored, preserved, safeguarded, and made accessible to users.

A **Digital Archiving Life Cycle Model** reflects the key characteristic of digital archiving, its circular and continual character, a cyclical process in which stages follow one after the other continuously, without an end point. The Life Cycle Model also draws attention to the need for taking actions and actively managing a digital archive throughout its life cycle. At the same time, the Model clearly presents the wide scope of responsibilities involved in the digital archiving process. Finally, the Life Cycle Model makes clear that decisions and actions in each phase affect what and how can be done in each subsequent stage and any new iteration of the process.

The Digital Archiving Life Cycle Model applied in this manual is developed to tailor to the specific needs and challenges of CSOs. It reflects some of the elements of the OAIS Reference model and partly the structure of the DCC Curation Life Cycle Model. The OAIS Model is the most widely used model for digital archiving, while the DCC Life Cycle Model includes many of the considerations that also affect CSOs engaged in digital archiving.

The **General Plan** of a digital archive is the first and crucial step in the process of its development. It lays out the reason and the method for the archive's development by providing it with Guiding Principles as well as key decisions regarding the content, access, and major organizational, technological, and resource-related issues. The widely scoped, detailed, and advanced planning contained in the General Plan will help the organization navigate a wide array of challenges that need to be met in the later stages of the process of digital archive creation.

The **Guiding Principles** of a digital archive summarize the reasoning behind its development. They state why an archive is needed, who will be using it and how, and what the expected benefits of its creation and development are. The Guiding Principles also address the required resources and technologies, legal and security-related responsibilities, and organizational matters.

The **Identification Inventory** is the initial list of item groups we can identify in the material we wish to digitally archive. It is a table that lists the identified item groups and includes information about their type, format, size, amount, condition, location, and storage space or storage media. The Identification Inventory provides us with a clear overview of what source material we have, in what quantity, and in which shapes and forms.

Selection and Prioritization are archival procedures in which we make decisions on which groups of source materials should be preserved and for how long, and what the order of their preservation should be.

Organization of material for archiving involves introducing a certain logical and hierarchical order into it and thereby devising its structure. This is done on the level of item groups identified through the Inventory, using the organization's knowledge and understanding of the material.

The process of organization of the selected material means the entire content of selected material is divided into several fundamental groups, each based on one or more common features shared by the item groups they contain. These most generic groups are then divided into smaller subgroups of material, and so on until the level of individual items.

A **Table of the Archive's Structure** is an advanced version of the identification Inventory, which reflects a hierarchical arrangement of series and collections of item groups. In the table of the archive's structure, the main units of analysis – described in terms of their size, format, amount, etc. – are not only individual item groups but also series and collections of these item groups, arranged

in a hierarchical way. The table of the archive's structure hence contains information about our material, metadata (on the archive's structure, grouping of files in collections, series, subseries, and folders), and additional descriptive and technical metadata that we selected to add into it. It is a necessary tool that allows proper archival processing of the source material and the archive's organization and management.

Description of Archival Material enables the archive's proper preservation and guides future users by providing important contextual information. The content of an archive needs to be described in a way that will allow anyone to search for, locate, and access items in the collection, thereby enabling connections to be established between items, even from different groups. Simply put, without description, an archive would be more like simple storage in which it would eventually become impossible to find or manage content.

A **Digital Archiving System** is a system of software and hardware components that consists of databases, software tools that manage databases, and storage media; it is the technological infrastructure of a digital archive. It defines the scope and limit of the archive's functions and is instrumental in achieving its aim and goals. The main purpose of digital archiving is to ensure that the invaluable content we are preserving remains unchanged and accessible long into the future through an adequate and sustainable technological framework for the digital archive.

The databases and software tools are merged together and make up the main software component of the system – a digital archiving software that allows us to manage an organized collection of information about the archival material. The digital archival material itself is located on storage media, which are typically different physical devices that store, retain, and make digital archival data available for retrieval. The software and hardware components of a Digital Archiving System work together to enable performance of the key functions of a digital archive. These include storing, backing up, preserving, maintaining integrity and authenticity, safeguarding, providing access, managing, and eventually migrating archival data.

Digitization is a process of creating digital copies, or "surrogates" of original physical items. These digital copies are then processed as digital archival objects. Different types of physical objects can be digitized and then stored on a variety of media. They can, for example, include text, photographs, drawings, maps, video, audio, and other types of content, stored on paper, audio cassettes, VHS tapes, or any other physical or analogue storage media. The digital copies could also include objects such as pieces of clothing, banners, personal belongings, etc. The type of material that needs to be digitized will determine the procedures, technologies, digital formats, and other elements of any concrete digitization process.

Digital Surrogates are digital copies of physical archival items that are processed, preserved, and made accessible as digital archival objects. Digital surrogates can originate from different types of physical objects – documents, maps, video, artifacts, etc. – and can be stored in different digital formats.

Digitization Workflow is a plan or scheme that should include all digitization actions and operations – starting with review and preparation of physical items and workspace to the completion of the workflow through storing the created digital surrogates and making backup copies. Each digitization project will have its unique workflow, its specific sequence of actions and operations. Although specific actions and their sequences are tailored to each concrete project, we can identify the key elements required in any digitization workflow: preparation, process scheduling, digitization, quality control, post-processing, and storing and backup.

The **Digital File Name** of a digital archival item serves the very important role of descriptor of that particular item, which should contain information that allows us to identify what the item is and what it contains so we can locate it in the archive and properly manage and preserve it.

Optical Text Recognition (OCR) Software allows for the creation of fully searchable documents from originally non-searchable image files. In essence, by running an OCR software on our scanned image of a document, we add a layer of text onto that image file so other software can read it, which makes the

document fully searchable. This is essential for making human rights archives more accessible and visible, which is often a key purpose of their digitization.

Master Files are the best-quality files we produce through digitization and are intended to be preserved long term without loss of any essential features. The number of master files we will create will depend on the content of the originals and the planned uses of the digital surrogate. In addition to master files, we can also produce a number of secondary files, often called "access" or "service" files. These files are created from the master file and optimized for the intended use (e.g., for web or for research).

Metadata refers to the descriptions of archival items. It is essential that relevant metadata is collected and attached to the digital archival items and stored in a structured form suitable for software processing. Without its attached metadata, digital archival material becomes meaningless and unusable, as we might be unable to find or identify it, understand what it is, its context, history, creator, or where it belongs in the archive. Digital archival items allow for a range of metadata to be collected, such as technical specifications of an archival digital file or information about its creation or any further digital action taken on it.

Data Documentation provides information about the context of our data, our digital archival content, often in a textual or other human-readable form. Data documentation supplements metadata and provides information that enables others to use the archival content. Given that data documentation is also "data about data," it could also be seen as a specific type of metadata – one that provides context and is recorded in a human-friendly format.

Validation of digital archival files is the process of establishing whether they really are what we think they are. Through file validation, we check whether the format of a file is proper and correct – whether it is valid. In this way, we determine whether a file conforms to the specific file format specification or standards a specific file format such as .jpg, .doc., or TIFF must follow.

Fixity, a crucial element in long-term preservation of files, as well as in maintaining their integrity, authenticity and usability, refers to a state of being unchanged or permanent. In essence, fixity allows us to determine whether a file has changed over time or been altered or corrupted, and to track and record any such changes.

Digital Forensics in digital archiving refers to a set of software and application-based techniques that allow us to access and investigate digital archival material in relation to its authenticity, accountability, and accessibility. This is especially relevant for older data storage formats or when working with digital material of unclear origin and features, especially when a history of the material and "chain of custody" have not been established. Digital forensics allow us to, for example, extract relevant metadata, access content archived in outdated digital formats, establish a chain of custody, detect data manipulation and forgery, and identify issues with data privacy.

Ingest is the process of transferring digital items into the digital archive, during which the data, its metadata, and data documentation are stored and mutually linked within the Digital Archiving System. A set of preservation actions needs to be applied to the digital content in preparation for ingest, as well as after it has been finished. This includes scanning the digital files and backup copies with antivirus software and checking each file's fixity, validity, and quality.

An **Archive Access Plan** guides both the decision-making and implementation related to access to the archival material. It provides balance between the goal of assuring as wide an access to an archive as possible and the responsibility to safeguard data and adhere to legal and ethical norms regarding data privacy, sensitivity, confidentiality, and copyrights. A well-considered and clear Archive Access Plan will help achieve that balance.

An **Access Levels Scheme** is a part of the Archive Access Plan and provides an overview of "who has access to what, and how" in the form of a table.

Considering that different modes of access might need to be provided for different user groups, to material with varying permitted level of access, an Access Levels Scheme helps avoid confusion and errors.

A **Digital Archive Security Plan** guides the devising of security procedures and their implementation. Such a plan should list and describe security-related obligations of the archive based on the material it contains, security-related functions the archive needs to perform, security-related actions that will be taken to ensure the functions are properly performed, and tools and technologies needed for its implementation.

System Protection is the first function of data security for a digital archive, because in order to protect the content – the data – we must first safeguard its repository. System protection includes safeguarding against system failures as well as protecting the Digital Archiving System from malicious acts of corruption or deletion.

Whether it relates to private, sensitive, confidential, or copyrighted data, Data Protection is the essence of our archive's security planning. There are three main instruments, or actions an archive can take to protect its data: access control and management, redaction, and encryption.

Active Maintenance of a digital archive is necessary both to keep the digital archive operational and also to ensure the long-term preservation, authenticity, and access to its digital content. If the format of the digital files becomes obsolete, or if the storage media fails or backup software is outdated and flawed, the archival digital content may be compromised, damaged, or lost altogether, along with all the work put into building and developing the digital archive. Active maintenance of a digital archive requires a systematic approach and regular performance of a set of actions that include monitoring and migration.

The **Maintenance Plan** is centered around the two main sets of maintenance actions. It lists, describes, and schedules the execution of maintenance

monitoring and migration activities. The specific elements of the Maintenance Plan, such as time periods for regular checkups or concrete procedures, are defined in line with the needs of a given archive.

Monitoring and Preservation of archival data are actions that need to be performed on the digital archival data in the maintenance phase. This is in essence a continuation of the work done as part of the preparation of the digital material for ingest, including the backup of data, checkups of file formats, validity, fixity, and quality assurance. Monitoring and preservation actions need to be planned and performed regularly to check and amend any irregularities or errors.

A **Disaster Recovery Plan** details how our data will be recovered or replaced in case of any major natural or human-caused failure, damage, theft, or malicious attack on our digital archival content or system. It is based on the existing backup arrangements that define the number of backup copies, their geographic location, and type of storage media used. The Disaster Recovery plan should provide instructions on which of our backup copies should be used in which disaster type circumstances and by which technological means to replace and recover the data.

Monitoring the digital archive software and improving it when needed ensures that it continues to meet our requirements and avoids it becoming obsolete. However, if the monitoring shows that a software can no longer be adapted, or that it is losing its support community, we will need to find a new appropriate software solution and migrate to it.

Monitoring Storage Media is necessary to timely detect any errors or damage, as well as to prevent it from becoming obsolete or outdated. Over time, storage media can become unstable and unreliable and cause data corruption or loss. A rule of thumb for a safe preservation practice is for storage media to have a short life cycle, sometimes estimated at only three to five years. This means

that after this period, we will need to find and obtain a new storage media and migrate our data to it.

The **Migration of Data, Software, and Storage Media** is an essential component of active maintenance of a digital archive, as it allows us to preserve our archival data by migrating it to new formats, software, or storage media. Migration needs to be performed timely, in a systematic and carefully planned way, following clear rules and including the set of preservation actions – checkups of fixity, validity, and quality assurance – as a mandatory step both before and after the actual transfer of files.

A **Resourcing and Fundraising Plan** should contain a thorough assessment of the overall resource needs related to the development of the digital archive, broken down by phases and activities. This should include an analysis of the more immediate short- and medium-term needs in developing the digital archive (i.e., which resources need to be obtained and within which time frame). Additional assessment should then be made of the resources the organization already has or can reasonably easily acquire. Based on these assessments, a realistic (time- and resources-wise) feasible plan should be made for securing the necessary resources and funds for developing a digital archive, both in terms of immediate steps and in the long term.

Outreach Strategy of a digital archive describes its outreach goals and beneficiaries and a plan of activities. It can help achieve the activism goals of the archive – be it to inform, educate, raise awareness, or to seek truth, accountability or reparations – and also to make them more feasible and attainable. An archive with a bigger presence and impact in the community, with wider visibility and credibility, will be more likely to engage a range of actors necessary for its long-term sustainability. This includes expert staff, users, CSOs, and other partners, donors, teachers, students, universities, supporters, mentors, journalists, and others who can and need to contribute to a digital archive's successful creation, maintenance, and development.

PART 14

LITERATURE AND RESOURCES

Application of International Standard Archival Description – ISAD(G) – for Human Rights Archives
Human Rights Working Group, International Council of Archives
25 October 2012
https://www.ica.org/sites/default/files/HRG_2012_toolkit_ISAD-G_EN.pdf

Caring for Collections – Preservation and Conservation
Resources from Emory University Libraries website
https://guides.libraries.emory.edu/c.php?g=50420&p=325032

Community Owned digital Preservation Tool Registry (COPTR)
https://coptr.digipres.org/index.php/Main_Page

DCC Curation Life Cycle Model
Sarah Higgins
INTERNATIONAL JOURNAL OF DIGITAL CURATION, VOL 3, No. 1 (2008)
http://www.ijdc.net/article/view/69/48
Digital Curation Centre website
https://www.dcc.ac.uk/guidance/curation-lifecycle-model

Digital Preservation Handbook
Digital Preservation Coalition
https://www.dpconline.org/handbook

Dublin Core Metadata Element Set (Dublin Core)
The Dublin Core Metadata Initiative
June 2012
https://www.dublincore.org/specifications/dublin-core/dces/

Levels of Born-Digital Access
Digital Library Federation – Born-Digital Archives Working Group
Council on Library and Information Resources, USA
February 2020
https://osf.io/hqmy4/download

Living Archives – A toolkit for CSOs in the Creation of Human Rights Oral Archives and Organizing Their Documentation
The International Coalition of Sites of Conscience
2020
https://gijtr.org/wp-content/uploads/2021/12/Living-Archives-Toolkit-6x9-EN-final.pdf

OAIS – Reference Model for an Open Archival Information System
International Organisation for Standardization and The Consultative Committee for Space Data Systems, June 2012
https://public.ccsds.org/pubs/650x0m2.pdf
OAIS website
http://www.oais.info/

PREMIS Data Dictionary for Preservation Metadata
Library of Congress, USA
November 2015
https://www.loc.gov/standards/premis/

Selecting Storage Media for Long-Term Preservation
The National Archives, United Kingdom
August, 2008
https://cdn.nationalarchives.gov.uk/documents/information-management/selecting-storage-media.pdf

---

Supporting CSOs in Digital Archiving
Resources webpage of the Global Initiative for Justice, Truth and Reconciliation
https://gijtr.org/supporting-csos-in-digital-archiving-1/

Technical Guidelines for Digitizing Cultural Heritage Materials
USA Federal Agencies Digital Guidelines Initiative
https://www.digitizationguidelines.gov/guidelines

UN Archives and Records Management Section Website
https://archives.un.org/content/digital-preservation

Web Content Accessibility Guidelines (WCAG)
Web Accessibility Initiative
https://www.w3.org/WAI/standards-guidelines/wcag

PART 15

ADDENDUMS

ADDENDUM I

FICTIONAL EXAMPLE OF A HUMAN RIGHTS ARCHIVE'S GUIDING PRINCIPLES

GUIDING PRINCIPLE 1: PURPOSE STATEMENT

Description

The first Guiding Principle needs to clearly communicate the main reasons why we are building the digital archive and what will the archive do. The purpose should reflect the organizational and community needs that give rise to the creation of the digital archive.

The Purpose Statement should be kept brief and to the point with one or two simple sentences.

Human rights CSOs typically cite two main purposes for building digital archives: preservation of invaluable material on human rights violations, and provision of wider or easier access to it. Other frequent reasons behind development of a digital archive include legal, educational, research, advocacy, memorialization, and other functions, often related to transitional justice processes.

Example

We collect, digitally preserve, and safeguard material related to human rights violations committed in Georgia in order to save them from loss and harm and make them publicly available in Georgia and worldwide.

GUIDING PRINCIPLE 2: AIMS, GOALS, AND BENEFICIARIES

Description

The Guiding Principle 2 defines in more concrete terms what the archive is designed to achieve and who its expected beneficiaries are.

The "aim" – which refers to a wider function the archive will provide for its community, specific groups, or society as a whole – is of the archive outline of the social benefits the archive is seeking to bring. There could be anywhere from just one to several aims of the archive, which should be clearly stated and defined.

The goals should specify the types of educational, academic, media, legal, or other projects and outcomes the archive aspires to support and enable. While it is important to state and outlie all of the archive's main goals, the list should not be excessive. The goals could also include examples of concrete products, events, or actions the archive aims to help produce and create.

The aims and goals of the archive are focused on one or several groups, organizations, or communities – the archive's beneficiaries. Beneficiaries include the archive's users and other groups that are positively affected by its work and the services it provides, such as victims and their families, legal professionals, teachers, students, journalists, and others. Organizations can also be beneficiaries (e.g., a truth commission relying on the archival content for its work, or other human rights CSOs).

Example

Aim: To facilitate transitional justice processes related to dealing with the legacy of mass human rights violations of USSR Police Forces in Georgia.

Goal 1: Promote access to and the use of the digital archive by journalists, students, teachers, researchers, artists, and others who may utilize it in their public-facing work.

Goal 2: Support use of the digital archive for development of outreach and educational tools and resources, as well as research, media, and artistic projects.

Goal 3: Develop an online platform that will streamline access to the digital archive for the general public and provide a space for publication and promotion of archive-based projects, actions, and products.

Beneficiaries

Main beneficiary groups include:

- Users of the archive
- Victims and their families
- State institutions dealing with the legacy of mass violence during the USSR period
- Other CSOs dealing with the legacy of mass violence during the USSR period
- Teachers
- Students
- Historians
- Other academic and non-academic researchers

GUIDING PRINCIPLE 3: SCOPE OF CONTENT

Description

Guiding Principle 3 determines what material will be included the archive, the defining of which allows us to focus time and resources on the items that are most relevant and important, including by guiding us in deciding which material should not be included in the digital archive. The scope of content should define what should be made part of any given archive based on criteria that are most relevant for its specific content and context.

Example

The following are the criteria that are usually important to be included in defining a digital archive's scope of content:

- The types and formats of items. (e.g., paper documents, photographs, and analogue multimedia material)
- The geographical area to which the items are related and time period (e.g., USSR Republic of Georgia from 1939 to 1989)
- The themes or topics the items are about (e.g., human rights violations by representatives of USSR police forces)
- Sensitivity, confidentiality, and credibility of material. (e.g., archive will include sensitive but not confidential material)

GUIDING PRINCIPLE 4: STATEMENT OF RESPONSIBILITIES

The Guiding Principle 4 defines the main responsibilities the digital archive has and was created to uphold. These predominantly, though not exclusively, include ethical-, legal-, access-, and data security-related duties.

These are all highly relevant issues for human rights CSOs that require the organization to regularly revisit the Statement of Responsibilities and further devise or revise it as necessary to ensure it remains comprehensive, up to date, and effective.

Note: *The following example lays out areas the Statement of Responsibilities should include. Depending on the content and context of a specific digital archive, the Statement should provide concrete considerations and measures the digital archive will take to uphold these responsibilities. As the archive develops over time, it will be possible to define its responsibilities in more concrete and detailed terms.*

Example

The digital archive will take all reasonable measures to ensure its work is aligned with the following responsibilities:

- Protection of data privacy
- Respect of intellectual property and copyrights
- Trustworthy and confidential management of sensitive data, or data that otherwise requires such approach
- Control over access and dissemination of the archive's content
- Safeguarding data from harm, loss, or unauthorized access
- To maintain credibility and integrity of data, including through control of the chain of custody.

GUIDING PRINCIPLE 5: IMPLEMENTATION STATEMENT

The Guiding Principle 5 responds to the question of "how" the other principles will be achieved and realized. The Implementation Statement should cover three main aspects of a successful and sustainable digital archive: resources, organization, and technology.

The Statement should deal with these domains in a general way, laying out the key considerations and approaches. More detailed, operational decisions will be devised through implementation plans, which will be developed throughout the process to address specific elements of implementation.

Without considering and maintaining each of these components a digital archive cannot be sustained. These three components need to be considered together to provide for a feasible development and long-term resilience of a digital archive. It is important that a balance is struck between technology, people, funding, and the organizational practices.

Example

Note: *This example lists the areas the Implementation Statement should include. Details on each of these domains need to be included based on the specific circumstances of the given archive.*

This digital archive operates based on advanced planning of the resources and requirements needed to develop, populate, and continuously maintain and develop it. The plans include cost assessment and a time frame in which these resources can be obtained and put into practice. Building this digital archive is supported by the entire organization: management, IT, fundraisers, project managers, field researchers, communications, etc. The roles and responsibilities are clearly defined and aligned with each other, aimed toward achieving the same joint purpose.

The technology elements of this digital archive include the hardware, software, and secure environments required to create, develop, manage, maintain, and

provide access to it. This digital archive will develop measures to ensure its continued readiness to adequately respond to the ever-changing technological developments.

Resources

- Types of resources needed include financial, human, organizational, and technical
- Estimate of requirements for each type of resources needed in initial phases of digital archive development for long-term maintenance
- Outline of modes and approaches through which the organization will sustainably obtain each type of the required resources, and over which time period
- For the resources that cannot be obtained, the organization states alternative ways of ensuring that the goals of the digital archiving process are achieved, either by identifying different solutions or redefining the goals that can be achieved.

Organization

- Commitment to developing timely, specific plans for key areas of the archive's development, maintenance, and continued transformation
- Establishment of procedures and workflows to ensure proper performance of key functions and upholding of the digital archive's responsibilities
- Documentation of key decisions and overall operations of the digital archive to create a reference and knowledge base for its future development
- Ensuring the necessary staff training and expertise to build the organization's capacities and the hiring of relevant external experts when necessary

- Identifying possible risks and then determining approaches to how they could be mitigated
- Clear division of roles and responsibilities within the organization and in relation to external actors

Technology

- Selection, installation, maintenance, and timely transformation of the Digital Archiving System
- Selection, installation, maintenance, and timely migration of the storage and back-up system
- Utilization of appropriate archiving software and tools, including for preservation actions
- Use of data safety and archive security tools and solutions to prevent leakage, harm, and loss of data
- Use of online technology, software, and tools needed to provide secure and appropriate access to different segments of the archive for the general public and individual users

ADDENDUM II

HOW TO ORGANIZE A PHYSICAL ARCHIVE IN 10 STEPS

Marc Drouin

Introduction

This guide presents a 10-step process for organizing documents related to human rights violations and transitional justice. It is intended for civil society organizations that have little to no archival experience or that have never tried to salvage and organize the documents they have accumulated over many years of work. Organizations that already have a physical or digital archive will have completed most, if not all, of the suggestions proposed here.

The 10 steps correspond to three basic themes: the first is very practical and corresponds to the staff, space, and materials needed to organize the documents (steps 1 through 5); the second is more theoretical and requires thinking about document categories and subcategories that will provide the final archive with a logical structure (steps 6 through 8); and the third section addresses creating an inventory describing each document in the archive and providing basic treatment for the conservation of each document (steps 9 and 10).

Since every archive is unique, the 10 steps are merely suggestions that can be followed to organize a basic physical archive prior to the digitization of documents, a phase that follows the process proposed below. If the 10 steps are followed and all the documents in the physical archive are organized, then digitization can be carried out in the same order. Therefore, it is important to understand that any document in a physical archive must be organized, described, and stored in suitable conditions before creating a digital version, whether it will be used for the organization's own purposes or for public disclosure.

Before implementing the 10 steps, we recommended assembling the following supplies, equipment, and staff:

- A team of two or three people responsible for carrying out the archival process;
- Personal protective equipment, such as gowns or oversized long-sleeved shirts, masks, latex or lint-free cotton gloves, hair nets or caps;
- Cleaning materials for the workspace;
- A dry, clean, well-lit, ventilated workspace;
- Worktables, chairs, and shelves that are painted or made of stainless steel;
- One or two fans (only if no spores are found on the documents);
- Cardboard boxes with lids for organizing and storing documents;
- Post-it labels/sheets of paper, tape, markers, and stainless-steel staples;
- Legal size manila folders;
- Reams of acid-free paper in letter, legal, and oversized formats;
- 1-centimeter wide twill tape (sold in rolls by the meter);
- Reams of chipboard;
- A paper cutter and/or scissors and Exacto knife;
- A metal ruler or T-square (in centimeters or inches);
- Plastic paper clips;
- Brush with soft, natural bristles (not plastic);
- Letter opener or lab spatula to remove staples (do not use staple removers);
- At least one workstation with a computer with Excel software.

STEP ONE: FORM A TEAM

From the moment a human rights organization makes the decision to create and maintain an institutional archive, a team will be charged with fulfilling this new goal. Because it includes sensitive information that can compromise the safety of victims, witnesses, and/or their families, the organization, protection, and safekeeping of the archive is a responsibility that goes beyond the mere preservation and administration of documents.

For this reason, the archive team may include members of the organization's board of directors who provide general guidance; staff that will carry out the organization and preservation of the documents; and trusted external consultants who provide specific technical and professional advice and skills as the archival process progresses. At the end of the document organization process, together, these individuals will ensure the maintenance and integrity of the archive in the longer term.

The team may vary in number depending on the available budget, the size of the organization, and the volume of documents to be organized. In some cases, it is possible to appoint one or two people to manage the process. In continual consultation with other members of the organization, the team will make decisions regarding archive content, purchasing materials, and hiring specialized personnel when necessary.

Organizations with limited resources can seek the support of other organizations with previous archival experience or set up agreements with academic institutions to have students work at the archive as interns. As the archive grows, the organization can train some of its members to carry out more specialized work for the administration and eventual digitization of its contents.

Initially, the team will lead a process to salvage and safeguard documents that, due to a lack of resources, staff or space, may have been previously stored in a makeshift or disorderly manner. The goal here is to build a team that has the technical capabilities and ethical responsibility to create and maintain the archive.

Goal: Create and Maintain an Archive (PHOTOS)

The archive team will have to make important decisions regarding the criteria for selecting materials to be preserved and archived. Although organizations often want to preserve all of their materials because they all seem valuable, it is often not reasonable or sustainable to do so , as some materials may be duplicated or do not provide relevant information about the organization, its history, or its activities. In addition, it might be too expensive or impractical, in terms of the space and equipment required for long-term storage, to keep all the materials. These are important decisions that must be made by the team in charge of organizing and maintaining the archive.

STEP TWO: **LOCATE THE MATERIALS**

Over the years, organizations produce a large number of documents in the course of their operations and activities. As the volume of documents grows, the materials are often stored in inappropriate places to create space for more current documents or for those that are used on a daily basis. In many organizations, older documents are tied into bundles or kept in bags or boxes and stored in empty rooms, closets or storage spaces where they are exposed

to moisture, heat, and pests. When an organization moves its offices, sometimes it takes these bundles of materials to its new location, distributes them among members of the organization, or gives them to other groups with the hope that one day they will be organized and preserved.

During this step, it is important to locate all the organization's scattered materials and set up an appropriate space for their centralization, organization, and preservation.

Locate and centralize the documents to be organized

STEP THREE: FINDING A SUITABLE WORKSPACE

The workspace must offer protection from the elements in order to avoid any further physical damage to the documents. Even though this space may not be the one where the archive is eventually stored, it is needed at this time to centralize the documents and begin to review and organize them. This space must be clean and dry, free of pets and pests, and have little traffic. This is where the initial organization of the documents will begin and the process of identification, cleaning, organization, and preservation will be carried out. Ideally, the space should be properly lit and ventilated, furnished with worktables and shelves, and have doors that allow for controlled access to the documents.

Work in a suitable space



In particularly arid, humid, or dusty environments, we recommend the workspace be equipped with hermetically sealed windows and doors. It is important to avoid dirt or carpeted floors because they can be very damaging to documents. When possible, the color of the floor should allow dust particles to be easily seen so that regular cleaning can be done more easily and effectively.

If the space permits, one or more fans can ensure adequate ventilation. However, the use of fans should be avoided if the documents contain spores or fungi, in which case the moving air could spread the spores around the space and thereby harm staff, the integrity of the documents, or both. If an air conditioner is used, it must have a mechanical filter that can be changed frequently. Finally, the consumption of food and drink in the workspace should be prohibited to prevent damage to documents during handling. In addition, food and drink remnants can attract insects that might further damage and degrade the documents.

Protect your workspace from the elements



STEP FOUR: GETTING THE DOCUMENTS OFF THE GROUND

From the start, it is very important that the documents and materials to be organized are not directly in contact with the floor or walls. This reduces the possibility of contact with moisture or dust. In the very short term, you can use prefabricated platforms, build makeshift platforms with bricks and wooden boards, or use moisture-resistant plastics to cover the floor where the documents will be collected. In the medium and long term, ideally, you should use painted or stainless-steel shelves to ensure the preservation of the documents, as materials such as wood can attract moisture and termites, which are very harmful to paper documents. However, what is important at this point of the process is to provide a first level of protection for the materials that previously may have been exposed to unsuitable storage conditions.

Elevate documents off the floor and away from walls

STEP FIVE: PROCURE BASIC MATERIALS

Once the work team has located a suitable space to gather the documents to be archived and the documents are elevated off the ground and moved away from the walls, basic materials and furniture can be acquired for their initial organization. Worktables, chairs, shelves, and cardboard or plastic boxes will be needed at this point to organize the documents.

For the team members in charge of the archive, it is also important to acquire and use personal protective equipment such as masks, smocks or oversized long-sleeved shirts, latex gloves, hair nets or caps, as well as cleaning products and insecticide for the workspace.

A document stored in humid environments and extreme temperatures is likely to be infested with fungus, mold, and/or pests that deteriorate the document itself and threaten the health of the archivists. Although the protection and conservation of the documents is important, the health and welfare of staff must be the highest priority before, during, and after any archival process.

Tables and boxes to organize the archival documents

Useful items for archival organization



Personal protective equipment for team members



STEP SIX: DEFINITION OF GENERAL ARCHIVAL CATEGORIES

During this step, team members will consider how best to group documents together according to **general classification categories**. These **general categories** will reflect the need for creating the archive in the first place, while respecting the original purpose of the documents and their use in the past.

Identifying **general categories** is key to successfully organizing the archive. **General categories** provide an initial structure for document organization and their location in the archive. These categories may not necessarily be the only ones that will be used, but they will get the process started.

GENERAL ARCHIVAL CATEGORIES

NAME OF THE GENERAL CATEGORY_____

TYPES OF DOCUMENTS AND MATERIALS FOR THIS GENERAL CATEGORY:

NUMBER OF BOXES FOR THIS CATEGORY:_____

SUBCATEGORIES:

NAME OF THE GENERAL CATEGORY_____

TYPES OF DOCUMENTS AND MATERIALS FOR THIS GENERAL CATEGORY:

NUMBER OF BOXES FOR THIS CATEGORY: _____

SUBCATEGORIES:

NAME OF THE GENERAL CATEGORY_____

TYPES OF DOCUMENTS AND MATERIALS FOR THIS GENERAL CATEGORY:

NUMBER OF BOXES FOR THIS CATEGORY: _____

SUBCATEGORIES:

The **general categories** can reflect an organization's **work areas** or divisions. Depending on the organization, it may have an "Administration and Coordination Area," a "Communications Area," a "Training and Education Area," a "Research Area," a "Legal Area," a "Project Area," etc.

As part of their past activities, each of these **work areas** or divisions created and used documents, and these are the documents that now need to be organized and preserved in the archive. For this task, the organization's structural chart indicating its different work areas can help staff visualize the organization of the archive.

Archive team members may find it useful to draw a structural chart for the organization beginning with when it was founded and show how this structure has changed over time. The **work areas** of an organization may provide the necessary criteria to identify and define the archive's **general categories**.

There are also other **general categories** that can be used to organize archival documents, but the examples that follow in this guide will refer to an organization's structure by **work areas** to identify its archive's **general categories**.

General categories group documents that have something in common. The point of this step is not to describe each document, but rather to give the collection of documents a preliminary order that facilitates the following steps in the archival process.

At this point, it is important that the archive team create an initial **list of general categories** and, in the next steps, notes the different types of documents that are part of each **general category**.

The worksheet illustrated here, for example, can be used to identify and list **general categories**, as well as the various materials included in them and their amounts. Once the different **general categories** of the archive have been defined, they will be compiled on a single list.

STEP SEVEN: INITIAL ORGANIZATION OF THE DOCUMENTS

In this step, the documents and other archival materials are reviewed in order to classify them. This initial classification will make use of the **general categories** defined in the previous step and the documents and other materials will be placed in cardboard or plastic boxes. At the end of this step, all the documents in the archive will be grouped into boxes according to their **general category**.

At this time, the bags will be opened and the bundles of documents untied to determine their content. For this purpose, the team will use large worktables or a clean, plastic-coated floor, and it will need a sufficient number of cardboard

or plastic boxes with lids and a space with shelves to store the boxes in an orderly manner.

It is important throughout this and all other steps of the archival process that team members wear personal protective equipment, particularly gloves and masks, when working with archival documents.

During this step, you will encounter various documents: copies of legal files, press releases, payrolls, lists of members, minutes of board meetings, books, account statements, etc. The idea here is to classify these materials according to each **work area** of the organization that created and used them. This way, the documents created and used by a **work area** constitute a **general category**. In archival terms, this general category then becomes a **documentary series**.

Process for Organizing Materials	*Final Archive*
General Categories →	**Documentary Series**

Paper documents are likely to be loose or stapled and may have been placed in manila or plastic folders, ring binders, etc. If they are grouped in any way, at this time, you should not individualize and describe each document stored in a folder or binder, but rather place the folder or binder in a **general category**.

At this point of the process, you will likely encounter both letterand legal-size sheets of paper, and it is also possible that you will encounter other materials that contain information, such as photographs, floppy disks, hard drives, memory cards, audio and video cassettes, posters, maps, recognition plaques, etc. If possible, organize these materials in the same way as the others in the appropriate **general category** (even though their later handling may be different). If these materials do not have a label or other indications regarding their content, they must be set aside and mechanical or computer equipment must be found that can read or reproduce the information they contain.

Review and classify documents

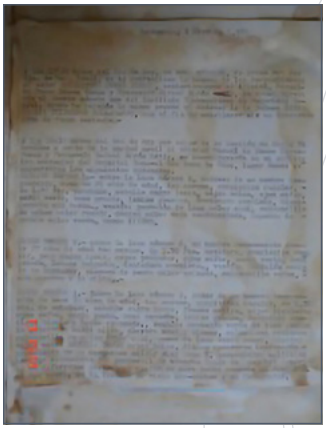
Once a document has been reviewed, it is placed in a labeled box. The label can be a sheet of paper taped onto the outside of the box indicating the following information: the name of the organization, the **general category** of the documents in the box, the consecutive number of the box in that category; and the **bookend years** of the documents (in other words, the year of the oldest document and the year of the most recent document in the box). You can also include a brief description of the contents of the box, such as "paper documents," "video cassettes," "floppy disks," etc., as well as the general condition of the materials.

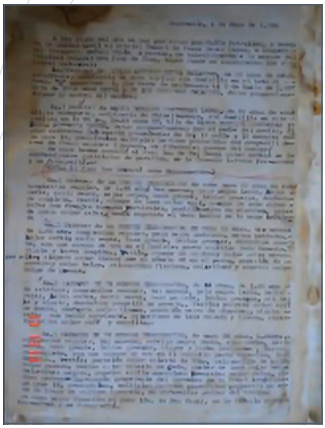
A box label can display this information as follows:

Organization name:	**Coastal Human Rights Committee**
General category:	**Administration and coordination**
Box number:	**1**
Time period:	**1979 et 1980**
Description of materials:	**Accounting binders**
Condition of the materials:	**Materials are damaged**

You might consider putting paper documents in one box and other types of materials, such as photographs, floppy disks or video cassettes in a separate box. Even though all these boxes belong to the same **general category**, the types of materials may be different. Another box must also be designated for materials in the same **general category** that are damaged or that require particular attention due to their condition.

Damaged documents





Storing documents in poor conditions can result in damage and loss of information. Therefore, every document in the archive must be handled with care, observing its physical condition and looking for any indication of damage. Such damage is usually biological or mechanical and, in general, is caused by light, humidity, and/or temperature. For example, light can discolor documents and weaken the fibers of the paper; humidity can generate spots and stains where dust and dirt accumulate; and high temperatures can encourage the growth of spores and fungi and attract insects.

Biological damage refers to living organisms such as rodents, insects, fungi, and bacteria that can destroy documents. Some of these, including cockroaches, worms, termites, and ants, consume the paper while breeding in a damp, dark

environment. Their presence can be identified by small holes appearing on the pages or by droppings left between the documents. Spores, fungi, and bacteria change the texture of the paper and weaken it. Their presence is evidenced by the appearance of black, reddish, or brown stains on the paper. A magnifying glass can be used to differentiate an acid stain from fungus., Fungi are superimposed on top of the paper and, in some cases, have thin bristles. .

Mechanical damage refers to the breakup of paper fibers or the impact of oxidation from metallic elements such as staples, clamps or clips. Also, the hardening or crystallization of adhesive tape used to repair a torn piece of paper can further tear or stain paper. Plastic elements, such as transparent or colored folders, damage documents due to their acidity and by encouraging humidity buildup. Mechanical damage is also caused by the tension on the paper due to bindings, clamps or presses, such as in classification binders. This tension is evidenced by tears, cracks, and the overall weakening of the paper.

Finally, damaged documents need particular attention. Temporarily, they should be kept in a box labeled "**Damaged Materials**." The label should also include other information about the **general category** to which they belong. Any indication of the presence of rodents, insects, fungi, or bacteria on or among the documents also indicates the need to regularly fumigate – in the absence of staff – the space where these documents are stored and handled.

After the documents have been reviewed and sorted into **general categories**, the archive team will have groups of boxes that correspond to the **general categories** of the archive. If possible, the boxes should be stored on stainless steel or painted wooden shelves. If shelves of this type are not available, the boxes can be temporarily stored on top of each other, taking care that no box is in direct contact with the ground and the weight of the upper boxes does not damage the lower boxes.

Before moving on to the next step, it is important that all archival materials are first classified into **general categories**. It is also important not to mark or write directly on the original documents. You should not repair tears with adhesive tape or add post-its or other similar items intended to highlight interesting information or documents. To preserve the integrity of the documents, we recommend keeping a record of the documents of interest in a notebook or other format that allows for their easy location and retrieval in the future.

In the list of **general categories** that was created in the previous step, the number of boxes that were filled for each **general category** is now indicated, as well as the total number of boxes for all categories. If, during the review and classification process, the team added or removed general categories, now is the time to update the list of **general categories** and the number of boxes that correspond to each.

Preliminary organization of the archive according to general categories

The shelves on which the archives are stored should be away from the walls and the documents should never touch the floor. This prevents any further damage to the documents from water leaks or any accidental spills in the storage space or facility.

STEP EIGHT: IDENTIFICATION OF DOCUMENTARY SUBCATEGORIES OR SUBSERIES

At the beginning of this step, the archive team can prioritize the **general categories** they want to inventory and preserve first. For example, due to legal considerations or ongoing investigations, the team may decide to work with the "Research Area" **general category** first, rather than documents from the "Administration and Coordination" work area.

Once the team has decided how to prioritize the **general categories** for processing, it will then move on to this step for the first **general category** and the contents of its respective boxes. This first **general category** of collected materials constitutes the first **documentary series** of the final archive.

Process for Organizing Materials	*Final Archive*
General Categories →	**Documentary Series**

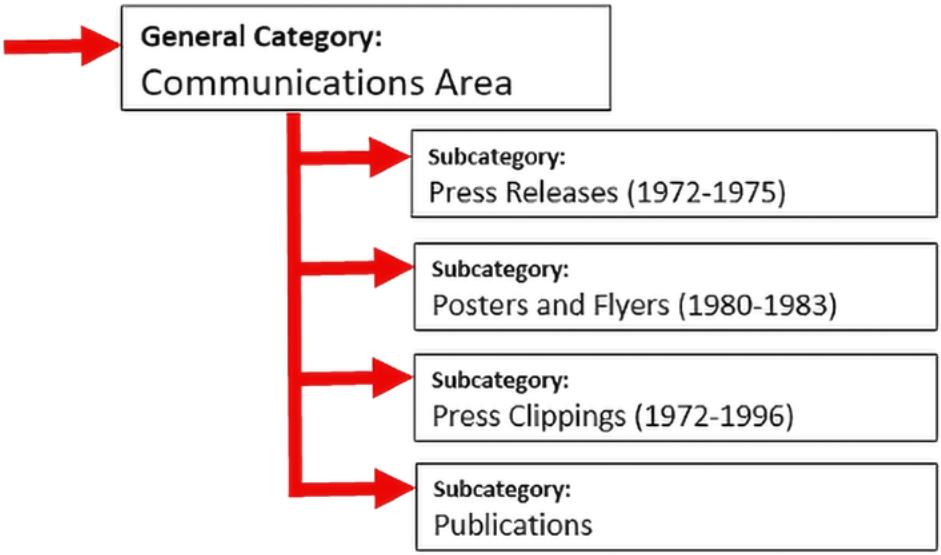
Since you will spend a fair amount of time working with and revising each document in the first **general category** or **documentary series**, it is important to have worktables, chairs, and fans —assuming there are no spores or fungi on the documents— to ensure good air circulation in the workspace. Tables, chairs, and the workspace should be cleaned frequently, and team members should use personal protective equipment (particularly gloves and masks) when handling documents. If smocks or loose-fitting shirts are used as cover-ups, they should be washed and dried frequently.

The goal of this step is to inspect and briefly describe all the collected documents and materials included in the first **general category** or **documentary series** and organize them into more specific **subcategories** or **subseries**. Subcategories generally refer to the types of documents contained in a **general category**.

For example, in the **general category** from the "Communications Area," you will find press releases, recordings of radio or television programs, photographs, press materials, pamphlets, flyers, posters, publications, video recordings, etc.

To better organize the various documents, these can be grouped into different **subcategories**, such as "Press Releases," within the **general category** "Communications Area." If the **general subcategory** is "Press Releases," then these can be organized chronologically, by year and by month, starting with the oldest releases going to the most recent; or thematically, according to the different topics of the releases organized in alphabetical order by topic, and then chronologically, by year and by month.

Taking the **general category** "Communications Area" as an example, the subcategories can be organized according to the different types of documents, such as "Press Releases," "Posters and Flyers," "Press Clippings," "Publications," etc. The relationship between the **general category** and its documentary **subcategories** can be illustrated as follows:



As the illustration indicates, all the documents related to the Communications Area are gathered in a single **general category** and the **subcategories**, identified according to document types or their function, allow them to be better organized in an orderly and logical manner.

What are subcategories?

When opening the boxes of materials from the first **general category**, the team will encounter various documents, such as those already mentioned, as well as different types of materials such as cassettes, photographs, posters, etc. Therefore, a **general category** can have one or more **secondary categories** based on the types of materials it contains. These can be organized into subcategories, with the number of **subcategories** depending on the diversity of the materials and how many **subcategories** the archive team deems necessary to organize a **general category**.

In this way, each **general category** or **documentary series** is made up of one or more **subcategories** or documentary **subseries**. This can be illustrated as follows:

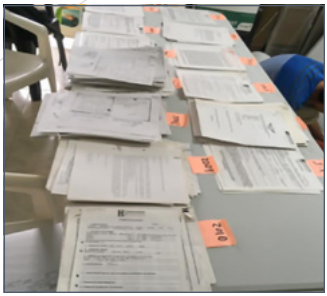
Process for Organizing Materials	*Final Archive*
General Category →	**Documentary Series**
• **Subcategory A** →	• **Documentary Subseries A**
• **Subcategory B** →	• **Documentary Subseries B**
• **Subcategory C** →	• **Documentary Subseries C**

Subcategories or **subseries** are smaller, more specific groupings of documents and materials. Not all **general categories** or **series** necessarily include **subcategories** or **subseries**, but their use allows for the more detailed organization of a **general category**.

To facilitate the identification of **subcategories**, it is important that team members pay attention to the types of documents they come across when they review them, noting the most striking characteristics of each. If the team identifies **subcategories** within a **general category** and decides to use them, it will place each of the documents of the **general category** into only one of the identified subcategories. Additional boxes can be used to separate documents and materials for each **subcategory**. A miscellaneous subcategory can also be provided for all documents in a **general category** that do not fall into one of the **subcategories** identified by the team.

Organization of general categories into subcategories





For example, within the **documentary series** "Training and Education Area," there may be **subseries** such as "Seminars," "Workshops," or "Training Camps." The "Seminars" **subseries** could be organized in chronological order for each seminar organized and carried out, from the oldest to the most recent, including documents such as announcements and other means of publicity for the event, participant lists, presentations, photographs, flipcharts, notes, minutes and reports, as well as budgets, tenders, and financial reports related to each event.

The **subseries** of the "Training and Education Area" **series** could also be organized according to the topics taught in the various training events, such as "Human Rights," "Indigenous Rights," "Children's Rights," etc. This **subseries** would contain documents for each activity in which that topic was taught, in chronological order, from oldest to most recent.

It is up to the archive team to determine the most appropriate **subseries** for each **documentary series** in the archive. The **series** and **subseries** are the main classification instruments for the organization of the final archive. At this point of the process, the **list of general categories** becomes a list of **documentary series** and the **subseries** identified by the team are added to the list.

STEP NINE: MAKING AN INVENTORY FOR EACH DOCUMENTARY SERIES AND SUBSERIES

This step consists of preparing and completing an **inventory form** (see example below) for the documents previously organized into **documentary series** and **subseries**. In this step, team members will again work directly with the **documentary series and subseries**, one at a time, in order to describe the contents of each document, its physical state of conservation, and its date of preparation to properly situate it in the archive. At this point, a specific and definitive place is assigned to each document within the physical archive.

On the **inventory form**, the characteristics of each document are indicated. The form can be filled out by hand, on a sheet of paper divided into **nine columns,** or directly in an Excel spreadsheet. If many people are going through the documents but few computers are available, then we recommend using paper forms to start with. Then someone from the team can enter the information from the forms onto a single Excel spreadsheet.

It doesn't matter whether the **inventory process** begins or ends with the Excel spreadsheet, but it is important to always keep at least two backup copies of the updated data. If paper forms are used, we recommend saving them, even after their contents have been transcribed onto an Excel spreadsheet.

The header of the **inventory form** indicates the name of the organization and the **documentary series** being processed. At this point, the inventory process returns to the **order of priority** of the **documentary series** and **subseries** that was determined previously.

As an example, the **nine columns** of the **inventory form** can include the following information:

1. The **number of the box** in which the document or material is stored;
2. The **folder** or **packet number** in which the document is located;
3. The **documentary series** to which the document belongs;
4. The **documentary subseries** to which the document belongs;
5. The **date** of preparation of the document;

6. The **type of document**, including the number of pages, if applicable;
7. A **brief description or summary** of the document's content;
8. The state of **preservation** of the document;
9. Additional **comments** on the document.

Example of a handwritten inventory form:

INVENTORY FORM

Historical Archive of the Coastal Human Rights Committee

Documentary series: Communications Area

Box	Folder	Documentary Series	Documentary Sub-series	Date	Document Type	Brief Description or Summary	Preservation Status	Comments
1	1	Communications Area	Press releases	19850711	Complaint, 1 page	Kidnapping of three farmers from the coast	Good condition	Original and a copy
1	1	Communications Area	Press releases	19850712	Declaration, 2 pages	Content of the meeting with the Attorney General	Good condition	Removed the rusty staple
1	1	Communications Area	Press releases	19850716	Advertisement, 1 page	Call for a demonstration in front of the Congress	Good condition, but oxidation spots from the staple	Removed the rusty staple
1	1	Communications Area	Press releases	19850714	Declaration, 2 pages	Declaration of the occupants of the Cathedral	Poor condition and missing second page	
2	2	Communications Area	Newsreel recordings	19850815	Video cassette, 14 min.	Channel 7 newscast, interview with director López	The volume is very low	Find out how to digitize the video
2	3	Communications Area	Fishermen Interviews	19850817	Audio cassette, 87 min.	Case of fishermen raiding the port	Bad, damaged by moisture	Inaudible

The same form in an Excel spreadsheet:

	A	B	C	D	E	F	G	H	I
1	**Box**	**Folder**	**Documentary Series**	**Documentary Sub-series**	**Date**	**Document Type**	**Brief Description or Summary**	**State of Preservation**	**Comments**
2	1	1	Communications Area	Press releases	19850711	Complaint, 1 page	Kidnapping of three farmers from the coast	Good condition	Original and a copy
3	1	1	Communications Area	Press releases	19850712	Declaration, 2 pages	Content of the meeting with the Attorney General	Good condition	Removed the rusty staple
4	1	1	Communications Area	Press releases	19850716	Advertisement, 1 page	Call for a demonstration in front of the Congress	Good condition, but oxidation spots from the staple	Removed the rusty staple
5	1	1	Communications Area	Press releases	19850714	Declaration, 2 pages	Declaration of the occupants of the Cathedral	Poor condition and missing second page	
6	2	2	Communications Area	newsreel recordings	19850815	Video cassette, 14 min.	Channel 7 newscast, interview with director López	The volume is very low	Find out how to digitize the video
7	2	3	Communications Area	Fishermen Interviews	19850817	Audio cassette, 87 min.	Case of fishermen raiding the port	Bad, damaged by moisture	Inaudible

Regarding the column "State of **Preservation**," it is simply a matter of making an intuitive assessment of the state of conservation of each document based on its appearance. Some specific detail can be given, such as "possible fungus," or it can simply refer to the general condition of the piece in terms of it being in "good condition" or "poor condition."

It is important at this point to give a definitive order to the **documentary series** and that each document of each **subseries** is organized according to the order decided on by the team, whether it is chronological, thematic, or alphabetical. To correctly locate a particular document in the future in its folder or packet and in its box, the final organization must be reflected in the **inventory**. Going forward, the **inventory** will serve as the primary tool to locate any document in the physical archive.

STEP 10: BASIC TREATMENT FOR THE PRESERVATION OF DOCUMENTS

As a counterpart to the previous step of inventory and the final organization of the archive into **documentary series** and **subseries**, in this tenth and last step of the process, documents are given basic treatment with the aim of ensuring their preservation and future digitization.

We recommend turning off fans during the document cleaning process so as not to spread dust and dirt from the documents, especially spores or fungi. As in the previous steps, staff responsible for carrying out treatment measures must ensure that they wear all the necessary protective equipment.

In this step, all metal elements prone to rust, such as staples, clamps, and clips are removed from the paper documents. Plastic elements, such as folders or acetate envelopes, should also be eliminated. Staples can be removed with a letter opener or lab spatula, always taking care not to tear the document and to properly dispose of recovered metal materials.

For the removal of staples in particular, the use of a staple remover is not recommended, since the pliers pull out the staples instead of carefully removing them, thus easily tearing the paper. With the lab spatula, unfold the two ends

of the staple on the back of the sheet, turn the sheet over, and lift the staple out without damaging the document.

Carefully remove all metallic elements from the documents



After removing the staples, clamps or clips, both sides of the pages of each document must be cleaned using a dry brush with soft, natural bristles, always sweeping from the inside out and frequently clearing away dust particles and dirt from worktables and floors.

Brush both sides of each page



Once clean, if you need to refasten the pages of a document, we recommend using a plastic clip (not metal) placed on top of a piece of acid-free 3 cm x 6 cm bond paper folded over the first and last pages of the document in such a way that the clip does not come into direct contact with the original. Although

stainless steel staples exist, they can be very expensive, and when documents are digitized in the future, it will be necessary to remove them again, once more compromising the integrity of the paper.

When each document of a **subseries** has been organized, they can be separated from the preceding and subsequent documents with a full sheet of acid-free bond paper. In this situation, a plastic clip to hold the pages of a document together is unnecessary, as the documents are separated by sheets of bond paper, one before the first page and one following the last page of each document.

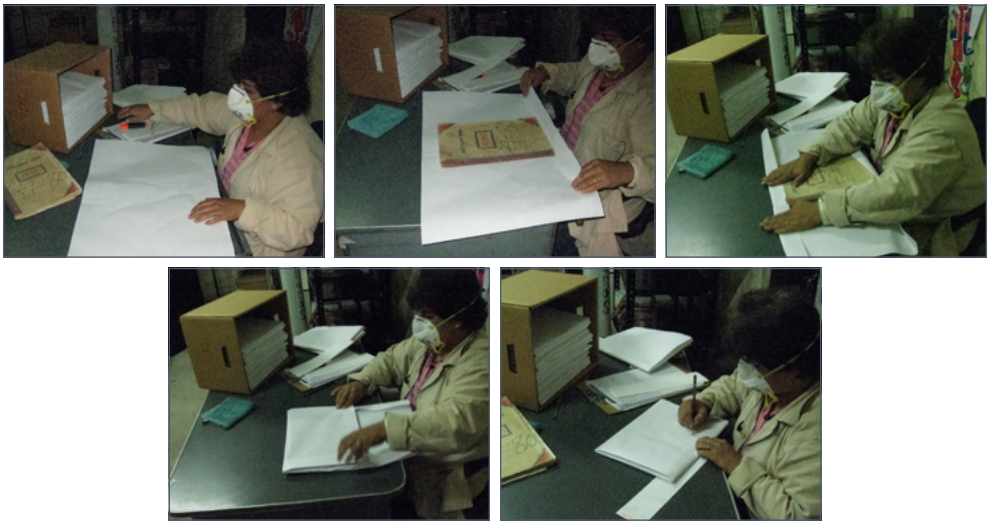
For example, multiple-page press releases from the "Press Releases" **subseries** of the "Communications Area" **series** can have a sheet of bond paper between each one. In this way, each press release is separated without using plastic clips. Multiple press releases, arranged in chronological order by date, can be bundled together in packets. You can place a piece of chipboard or posterboard on top of the packet and another below, then each packet is tied with one-centimeter-wide twill tape.

Create packets of documents organized chronologically





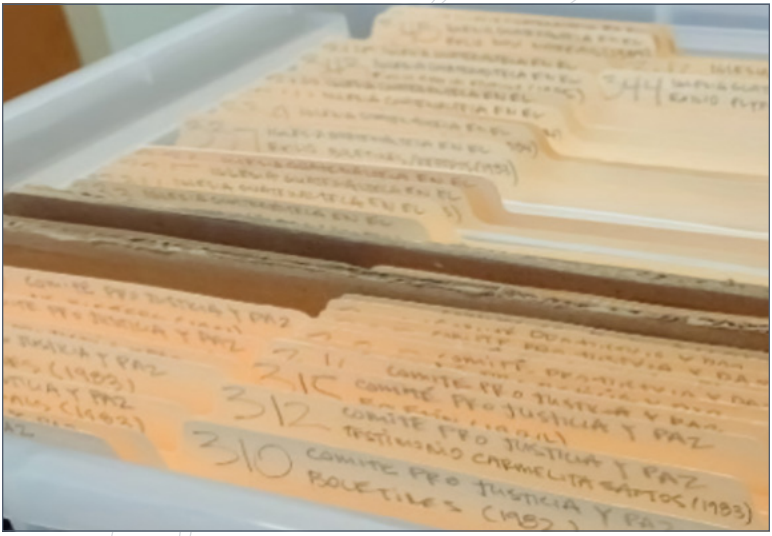
Materials other than stapled paper documents, such as notebooks, logs or newspapers, can be packed using acid-free sheets of paper, as illustrated below:



The same procedure can be applied to materials such as photographs, using chipboard or posterboard between each one and then packing several photographs inside a folded sheet of acid-free paper. Another alternative, in the event that the documents and materials are still being consulted or used on a regular basis, is to use legal-size manila folders to group documents in the same **subseries**.

Note: all packets, packages, and manila folders must be individually identified in pencil and their characteristics and contents recorded on the **inventory form**. Also, the sheets of bond paper used to separate documents and the chipboard or posterboard used to protect the packets of documents tied with twill tape must be the same size as the original documents.

Identify each folder with a sequential number that is recorded on the inventory form



As mentioned in step seven, all handwritten notes related to the documents should appear not on the original documents, but on the sheets of acid-free bond paper or on the archival folders that protect the originals.

At the end of the process, all the documents should be kept in boxes with lids and identified with a label that indicates the name of the archival collection, the documentary series, the documentary subseries, the consecutive number of the box, the bookend years of the documents and materials in the box, a brief description of their content, and their general state of preservation.

Below is an example of the information needed on a box's final label:

Coastal Human Rights Committee	
Documentary Series:	**Communications Area**
Documentary Subseries:	**Press Releases**
Box Number:	**1 of 5**
Time Period:	**1982-1984**
Description of Materials:	**Packets organized by date**
Status of Materials:	**Ready for digitization**

After the 10 steps have been completed, the organization's physical archive should look similar to the photograph to the left.

Finally, to the extent possible, the ambient conditions inside the physical archive room should be kept stable, controlling humidity, direct light, temperature, and dust.

The ideal temperature for paper is between 15° and 20°C. The ideal relative humidity should be between 45% and 60%, with a maximum daily fluctuation of 5%. The use of a hygrometer is recommended to conduct periodic measurements of the humidity.

It is also important to ensure that there are no sudden variations in temperature, as this can accelerate damage to archival documents and materials. If an air conditioner is installed, for example, it must be kept on 24/7 at the same temperature, as turning the unit on and off can cause abrupt changes in temperature that may accelerate document deterioration.

ADDENDUM III

EXAMPLE OF A LIST OF REQUIREMENTS FOR A DIGITAL ARCHIVING SYSTEM FOR A CSO ARCHIVE

A Digital Archiving System needs to meet the following requirements to be selected.

To perform to a high standard, it must include the following archival functions:

- Appraisal
- Description
- Arrangement
- Capturing and storing metadata
- Ingest
- Retention
- Backup
- Long-term preservation
- Maintaining integrity and credibility of data
- Recording chain of custody
- Safeguarding data
- Providing access
- Managing and administering data
- Migrating data to other formats, software and storage media

In addition to high quality of performance the Digital Archiving System is required to have integrated standards for these archival functions (i.e., the OAIS Reference Model and the standards related to it). In addition:

- The software and hardware components of the Digital Archiving System need to be compatible with each other, as well as with other technologies and information systems the organization already uses (e.g., databases and servers it needs to interact with).
- A Digital Archiving System is required to enable storage and management of all types and formats of data contained in the archive.
- The requirement is to minimize the costs related to purchase and maintenance of the Digital Archiving System.
- In terms of human resources, the requirement is that the Digital Archiving System can be implemented with no or minimal increase to the organization's staff. The Digital Archiving System should also provide opportunities for training of staff members working on the digital archive.
- The Digital Archiving System should provide strong user support and the possibility for quick access to expert assistance and troubleshooting.
- The longevity requirement for the Digital Archiving System is to provide for as long-term a solution as possible and feasible. The requirement is also to minimize any major maintenance, updates, or upgrades that are needed over a product's life cycle. A product that has a proven track record and a wide user community will be considered more likely to provide a long-term solution than will a novel product.
- Both software and hardware components need to be well-tested and have good to excellent user feedback, with proven reliable and consistent performance.
- The Digital Archiving System will be required to allow for streamlined, safe, and accurate migration and transfer of data to future file format, software, and storage media solutions.
- To facilitate efficient and streamlined use of the archival material, software that provides for an ergonomic and user-friendly interface is needed.
- The storage media must provide sufficient capacity for both current and future expected additions of material to the digital archive. This includes assessment of whether the storage media capacity can be easily and feasibly increased to provide for future needs.

- The selected Digital Archiving System must provide for adequate, regular, and sufficient backup of data. This includes storing the data in two or more separate storage units at separate geographical locations and, if possible, on different storage media.
- A requirement for the Digital Archiving System is that it provide for stable and safe storage (i.e., without changes or deterioration of records and ensuring that the data cannot be lost). The storage media should provide for a method of detecting errors in writing/reading of data onto storage, as well as for data recovery in case of failure.
- The storage media must be resilient to physical damage and other factors that can damage the records.
- Financial, technical, and human resources needed for obtaining, implementing, and maintaining the Digital Archiving System need to be feasible and sustainable for the organization in the long term.

ADDENDUM IV

RECOMMENDATIONS ON THE USE OF OPTICAL CHARACTER RECOGNITION SOFTWARE IN DIGITIZATION

1. Using the Right Software: Should be highly accurate, reliable, and work with multiple languages.

2. Using the Right Scan Parameters: When scanning documents, it is important to set the right parameters in your scanner settings. The foremost of these is orientation. Ensure that the document is fed into the scanner at the correct angle because a skewed scan can seriously affect Optical Character Recognition Software (OCR) accuracy. Test and tweak the settings until you achieve the desired result.

3. **Resolution Setting:** The best resolution for accurate OCR is 300 dpi. This level of resolution enables the OCR engine to work with twice the number of reference points versus 150 dpi.

4. **Color Mode Selection:** For discolored or old documents, RGB is the recommended color mode to enable the scanner to fully capture the contents of the physical document. In general, however, scanning in grayscale mode is the best option for OCR accuracy. Although the black and white mode helps the image be scanned at a faster rate, this could affect the quality of text recognition.

5. **Brightness and Contrast Adjustments:** For brightness, both extremes – too high and too low – can negatively affect OCR quality and accuracy. For that reason, 50% is the recommended brightness setting. However, this is also dependent on the scanner itself, so an initial trial and error phase may be expected. In terms of contrast, the highest setting is usually preferred.

6. **Image Correction and Decontamination:** These two components greatly impact the quality of OCR scanning. Image correction covers aspects such as increasing the resolution, applying color corrections, and trying out different contrast settings; decontamination involves the removal of non-text characters such as icons, non-text images, unusual characters, and so on. Both are important because they enable the OCR engine to "read" the document more accurately.

7. **Careful Manual Proofreading:** Depending on how accurate you want the end result to be, manual proofreading may be required – if accuracy is paramount, this is an indispensable step. It essentially involves human verification of a sample of processed files to ensure that the scanned characters are recognized correctly. It's a tedious and painstaking process, but essential in many cases.

ADDENDUM V

CONTRIBUTING CIVIL SOCIETY ORGANIZATIONS

GIJTR would like to acknowledge and thank all the CSOs that participated in the first research and exchange portion of this project, without whose insights the creation of this manual would not have been possible:

Afghanistan Human Rights and Democracy Organization (Afghanistan), National Coordination of Families of the Disappeared (Algeria), Área de Memoria y Derechos Humanos de la Fundación para la Democracia, Centro de Estudios Legales y Sociales, Memoria Abierta (Argentina), Belarusian Documentation Center (Belarus), The Post-Conflict Research Center, Udruženje za društvena istraživanja I komunikacije, War Childhood Museum (Bosnia and Herzegovina), Memorial da Resistência de São Paulo, Núcleo de Preservação da Memória Política (Brazil), Kdei Karuna, Legal Documentation Centre, Tuol Sleng Genocide Museum, Youth of Peace Organization (Cambodia), Contra Nocendi (Cameroon), Corporación Parque por la Paz Villa Grimaldi, Museo de la Memoria y los Derechos Humanos (Chile), Comisión Colombiana de Juristas, Comisión Intereclesial de Justicia y Paz, Consejo Comunitario General del San Juan, Movice Valle del Cauca, Observatorio Surcolombiano de Derechos Humanos, Pueblo Misak (Colombia), Documenta – Centre for dealing with the past (Croatia), Arab Reform Initiative (Egypt), Museo de la Palabra y la Imagen (El Salvador), Institute for Development of Freedom of Information (Georgia), Archivo Histórico de la Policía Nacional, Asociación Comunitaria Organizada de la Población Desarraigada en la Región Ixil, Asociación Familiares de Detenidos-Desaparecidos de Guatemala, Asociación Movimiento de Victimas para el Desarrollo Integral del Norte de Quiché, Asociación Q´anil Maya Kaqchikel, Centro de la Memoria Monseñor Gerardi de la Oficina de Derechos Humano del Arzobispado de Guatemala, Conferencia Nacional de Ministros de la Espiritualidad Maya de Guatemala, Coordinadora Nacional de Viudas de Guatemala, Grupo de Apoyo Mutuo, Memorial para la Concordia (Guatemala), Association Des Victimes Parents et Amis du 28 septembre 2009, L'organisation guinéenne de défense des droits de l'homme et du citoyen, Le Consortium des Associations des Jeunes pour la Défense des Victimes de Violences en Guinée (Guinea), Yazda (Iraq), The Women's Active Museum on War and Peace (Japan),

MAAN (Lebanon), Veritas (Serbia), Remembering the Ones We Lost (South Sudan), Syrian Center for Media and Freedom of Expression (Syria), Hafiza Merkezi, Syrian Network for Human Rights (Turkey), Center for Research on the Liberation Movement (Ukraine), National Security Archive, GW University (United States), Resonate!Yemen, The Abducted Mothers Association, SAM Organization for Rights and Freedoms (Yemen).

Endnotes

1 Please see Addendum IV for a complete list of the CSOs that contributed to the research phase of the project.

2 The key terms and concepts are listed in the order they are mentioned or referenced in the manual.

International Coalition of
SITES *of* CONSCIENCE

www.sitesofconscience.org

Facebook.com/SitesofConscience
@SitesConscience
SitesofConscience

GIJTR
Global Initiative for Justice
Truth & Reconciliation

www.gijtr.org

@GIJTR
@GIJTR