The Hala Protocol

for the Collection, Processing, and Transfer of Audio Data

The Hala Protocol is guidance for civil society to maximize the evidentiary value of audio data. Rooted in international standards, it breaks down legal, technical, and ethical standards into accessible best practices.

Last updated 5th August 2025

Hala Systems

Federal Foreign Office

The Hala Protocol

for the Collection, Processing, and Transfer of Audio Data

Coordinator

Ashley Jordana

Senior Researchers

Dr. Emma Irving

Sabrina Rewald, JD

Supporting Researchers

Petroula Alkistis Anastasiou

Uzay Yasar Aysev

Yana Ballod, Esq. (UA), LLM

Julia Freytag, LLM

Dr. Diletta Marchesi

Scott Martin

Advisors

Jason Antley
Sarah Bafadhel
Deniz Dirisu, LLM
Stefanie Frease
Lindsay Freeman
Kate Gibson
Dr. Jonathan W. Hak KC
Nerma Jelacic
Kate Keator
Dr. Alexa Koenig
Carlota Maldonado Montserrat, LLM
Dr. Ritumbra Manuvie
Caline Matar
Dr. Brianne McGonigle Leyh
Raji Abdul Salam
Basile Simon
Shannon Raj Singh
Emily Tripp
Raquel Vazquez Llorente
Mirjana Vukajlović
Dr. Sarah Zarmsky
Chelsea Zender Neely

The Hala Protocol is available online at **https://protocol.halasystems.com.**

The Hala Protocol

for the Collection, Processing, and Transfer of Audio Data

FOREWORD

Today, technology profoundly transforms how conflicts are chronicled, experienced, and understood. The human voice persists as a uniquely potent witness to history's critical moments. Unlike other forms of information, audio can tell unfiltered stories with unmatched immediacy and authenticity—commanders issuing orders, soldiers reporting atrocities, communications revealing strategic troop movements, and conversations that expose both intent and knowledge.

The power of audio has been demonstrated repeatedly across decades—from exposing the Watergate scandal to establishing responsibility for the downing of Flight MH17. These recordings penetrate even the most protected corridors of power, serving dual purposes: safeguarding legitimate state interests while ensuring accountability for those who abuse authority. In international courts from The Hague to Arusha, these ephemeral transmissions, when properly collected and preserved, have amplified victims' voices and brought perpetrators to justice.

Until now, a significant gap has existed in how we approach this vital evidentiary category. While comprehensive guidance exists for handling photographs, videos, and witness testimony, the unique challenges of using audio as evidence have remained largely unaddressed due to its complex technical nature. The challenges of voice identification, the privacy implications of intercepted communications, and the sheer volume of potentially relevant material all create distinct demands on those who have an interest in collecting and disseminating audio information. The Hala Protocol for the Collection, Processing, and Transfer of Audio Data (The Hala Protocol) fills this critical void. The Hala Protocol, developed by international legal experts and technologists from both academic and practical backgrounds, provides a comprehensive framework organized around the collection, processing, transfer, and ethical use of audio files. The 29 Best Practices therein aim to transform ephemeral sounds into enduring instruments of truth and accountability.

For those handling audio files—whether investigators, national authorities, journalists, humanitarian workers, peacekeepers, and civil society organizations alike—having clear, legally-grounded practices can mean the difference between justice and impunity. The digital realm is inherently fragile: transmissions vanish into the ether, files are deleted, and platforms change policies. In this context, the custodians of audio files bear a profound responsibility. By following the Hala Protocol Best Practices, actors will be better equipped to preserve crucial evidence with a greater chance of withstanding the rigorous demands of courtroom scrutiny.

The Hala Protocol transcends its role as a technical guide—it embodies a commitment that even amid warfare's chaos, the opportunity for accountability must endure. By transforming transient moments into lasting pieces of evidence, those who follow these practices contribute to a world where justice remains possible even in humanity's darkest moments—where the truth, properly preserved, can ultimately prevail against the brutality of conflict and impunity.

Sincerely,

John Jaeger
Chief Executive Officer

Ashley Jordana
Director of Law, Policy and Human Rights

INTRODUCTION

At the turn of the 20th century, warring parties were introduced to an array of new technologies that would help them communicate with one another during conflict. These methods of communication have evolved quickly throughout the ages, advancing the way that warring parties communicate with one another and the coordination of their efforts in theater. More recently, audio information flowing from these channels has been collected and used to document war crimes and crimes against humanity by providing real-time on-the-ground information that can help establish the truth of specific incidents.

Over the past decade, active conflicts have brought the importance of audio data for accountability into sharp focus. Following the Russian invasion of Ukraine in February 2022, members of the international community quickly noticed that Russian forces were using unencrypted radio channels to communicate. By monitoring and analyzing these communications, they were able to corroborate allegations of widespread war crimes by Russian soldiers.[1]

In the Netherlands, telephone communications helped aid the conviction of three accused in connection with the downing of flight MH17.[2] At the European Court of Human Rights, audio communications were used to support the allegations that Russia violated human rights on the territory of Ukraine.[3] Audio data has also played a prominent role at the International Criminal Court (ICC). In the 2021 *Ongwen* judgment, intercepted radio communications were used to establish the dynamics of the armed group, including the role of the defendant within the group hierarchy and how and when attacks were reported to commanders.[4] In the 2024 judgment in *Al-Hassan*, audio recordings helped show that the armed group intended to establish an Islamic State on the entire territory of Mali, which was an essential element of proving crimes against humanity.[5]

Civil society organisations (CSOs) are critical stakeholders in the accountability landscape and contribute substantially to the documentation of international crimes. Across recent conflicts, the documentation efforts of CSOs have helped ensure that audio data, which is ephemeral and highly volatile by nature, is preserved for accountability. A radio transmission or telephone call, if not intercepted and recorded in the moment, cannot be accessed later; a voice recording posted to social media can be removed by the platform or uploader without notice; and a voice note sent over an instant messaging app can be deleted by the sender within a certain time. As a result of this volatility, the copy of the data held by the CSO may be the only one in existence.

In recent years, a number of resources have been developed to help CSOs undertake documentation work effectively and responsibly. Examples include the Berkeley Protocol on Digital Open-Source Investigations, the GLAN/Bellingcat Methodology for Online Open Source Investigations, the WITNESS Video as Evidence guide, and the PILPG Handbook on Civil Society Documentation of Serious Human Rights Violations.

---

[1] Robin Stein, Christiaan Triebert, Natalie Reneau, Aleksandra Koroleva and Drew Jordan, 'Under Fire, Out of Fuel: What Intercepted Russian Radio Chatter Reveals' (The New York Times, 23 March 2022); Tom Hannen, 'Ukraine's battle of the airwaves' (Financial Times, 8 April 2022)

[2] Summary of the day in court: 17 November 2022 – Judgment (De Rechtspraak, 17 November 2022).

[3] *Ukraine and The Netherlands v Russia*, Decision, ECtHR, 8019/16, 43800/14 and 28525/20, 30 November 2022

[4] See *Prosecutor v Ongwen*, Trial Judgment, ICC-02/04-01/15-1762-Red, 4 February 2021 (*Prosecutor v Ongwen,* Trial Judgment), paras 858, 884, 1047, 1071, 1075, and 1079. See also Diletta Marchesi, 'Intercepted Communications in the *Ongwen* Case: Lessons to Learn on Documentary Evidence at the ICC' (2021) International Criminal Law Review 1.

[5] *Prosecutor v Al-Hassan,* Judgment, ICC-01/12-01/18-2594-Red, 26 June 2024, paras 436 (fn 1095) and 1288.

However, these standard setting initiatives do not address audio data specifically, resulting in an information gap surrounding how to best collect this type of data.

Audio data as a category of potential evidence of international crimes is often joined together with other categories of digital information, in particular video. While both forms of information deliver digital depictions of a moment in time surrounding the commission of a significant event, there are critical aspects of audio data that distinguish it from other forms of digital information. One aspect relates to its collection. First, audio data can be collected in bulk. Telephone and radio communications, as well as online audio data, has the potential to be collected on a large scale. Second, audio recordings can often be made without the knowledge (or consent) of one or more of the speakers. These two features create a situation that merits specific attention: a voluminous prevalence of audio (potential) evidence together with a heightened risk of violating privacy.

Furthermore, despite often being discussed alongside (or subsumed within) video material, case law from international criminal courts and tribunals suggests that video and audio are not treated identically. In particular, audio is less likely than video to be used as standalone evidence. For example, in the *Bemba et al* case, the ICC Trial Chamber elaborated extensively on how they could establish that the voices on the telephone recordings were those of the accused.[6] The Chamber did not rely solely on their own recognition of the voices, but rather looked at a number of other factors and corroborative evidence. This conforms with the approach taken at the ICTY, whereby intercepts were considered to be 'a special category of evidence in that in and of themselves, they bear no prima facie indicia of authenticity or reliability, and as such these requirements must generally be fulfilled by hearing from the relevant intercept operators or the participants in the intercepted conversation'.[7] Video, in contrast, has been relied upon as a standalone form of evidence.[8] This supports the proposition that it is harder to identify someone from their voice as compared with their (moving) image; and as a result, more evidence may be needed to corroborate standalone audio than video and photos.

The aim of the Hala Protocol is to help groups documenting human rights violations and international crimes, namely CSOs, maximise the evidentiary value of their audio data and improve the chances of the data being admissible in criminal proceedings. The Protocol provides 29 Best Practices for collecting, processing, and transferring audio data that are grounded in international law and case law. The Best Practices are designed to be incorporated into a CSO's workflow and align their data practices with evidentiary standards.

---

[6] *Prosecutor v Bemba et al.*, Judgment pursuant to Article 74 of the Statute, ICC-01/05-01/13-1989-Red, 19 October 2016, para 261*:* 'The Chamber did not rely on voice recognition alone to identify the speakers in a telephone conversation, but always considered the voices heard in connection with the call content and other relevant information'.

[7] *Prosecutor v Tolimir*, Decision on Prosecution's Motion for Admission of 28 Intercepts from the Bar Table, ICTY Case No. IT-05-99/2-T, 20 January 2012, para 14; see also *Prosecutor v Karadžić*, Decision on the Prosecution's First Motion for Judicial Notice of Documentary Evidence Related to the Sarajevo Component, ICTY Case no. IT-95-5/18-T, 31 March 2010, para 9: 'The Chamber considers intercepts to be a special category of evidence given that they bear no indicia of authenticity or reliability on their face [...] the authenticity and reliability of intercepts is established by further evidence, such as hearing from the relevant intercept operators or the participants in the intercepted conversation themselves.'

[8] In *Lubanga*, video was relied upon to establish that children clearly below the age of 15 were enlisted in the armed group: *Prosecutor v Lubanga,* Judgment Pursuant to Article 74 of the Statute, ICC-01/04-01/06-2842, 14 March 2012. See, for example: para 257 ('However, the video material, to a significant extent, "speaks for itself" and it falls therefore (along with the account of the witness as regards its content) into a separate category.'); para 711 (The Chamber has independently assessed the ages of the children identified in the video footage, to the extent that it is possible to draw a safe conclusion based on their appearance.); para 1262 ('On the basis, in particular, of the video footage the Chamber is persuaded there were children below the age of 15 who were responsible for ensuring the security of the accused during public events.'). Further, see discussions by the Chamber of a video in paras 792 and 793. In one instance, a video was relied upon even though the witness was discredited (para 268).

Given the ephemeral and volatile nature of audio data, and given that a CSO may hold the only copy of the data, working in alignment with evidentiary practices is crucially important. Following the Hala Protocol Best Practices can help CSOs to make their data more reliable and of greater potential probative value, enhancing its significance for accountability.

Audience

The CSOs that will benefit most from the Hala Protocol are those working to leverage audio data for international criminal accountability. The term 'civil society organisation'[9] is intended to be understood broadly, encompassing formally incorporated organisations, as well as informal groups such as university-based investigation labs,[10] which are now a prominent part of the digital investigations landscape.

Government actors, judicial institutions, and intergovernmental fact finding bodies are not direct addressees of the Protocol, yet they may derive benefit and guidance from it. This is due to the fact that these actors operate within existing ethical and legal frameworks that authorise, guide, and limit the scope of the actions they can take. Best practices applicable to these actors should therefore be tailored to such frameworks, rather than the broadly applicable approach of the Hala Protocol. (Inter)Governmental actors and judicial institutions may be authorised to take actions that non-governmental and non-judicial institutions are not, and may also face restrictions that non-governmental and non-judicial institutions do not.

Actors working outside the criminal justice context may also derive benefit from the Protocol. Journalists and advocacy organisations, for example, may choose to implement the Protocol in their information gathering processes to help bolster their work against allegations of mis- and dis-information (and in anticipation of their work potentially being useful in formal accountability at a later time).

The Role of Hala Systems Inc.

Hala Systems is a humanitarian technology company that develops solutions for civilian and asset protection, accountability, and the prevention of violence. The Hala Protocol was developed in the context of this work and builds on Hala's experience working with evidence collection in conflict zones and accountability for international crimes. The Hala Protocol was developed using public funding and is a public facing, open-source document available at no cost to all users.

The Objectives of the Hala Protocol: What is it for?

The Hala Protocol is structured into two parts. Part 1 contains the Best Practices, which are made up of an action statement, an explanatory note elaborating on the action statement, a technical specifications and

[9] The term civil society organisation (CSO) is preferred over the term non-governmental organisation (NGO) because the former encompasses the latter while still including other formations. See PILPG, Handbook on Civil Society Documentation of Serious Human Rights Violations: Principles and Best Practices (2016), page 18 and UNDP, Working with Civil Society in Foreign Aid, Possibilities for South-South Cooperation: NGOS, and SCOS: a note on terminology (3 November 2015), ANNEX I.

[10] See, for example, Utrecht University's Open-Source Global Justice Investigations Lab, Essex University's Digital Verification Unit, and UC Berkeley's Human Rights Investigations Lab.

resources section (where applicable), and a condensed legal framework section. The condensed legal framework contains links to the principal Legal Framework, which is found in Part 2 of the Protocol.

The different sections of the Best Practices reflect the three core objectives of the Hala Protocol.

1. **The first objective of the Protocol is to help CSOs maximise the evidentiary value of audio data**. This objective is achieved by the action statement and explanatory note of each Best Practice. These sections contain actionable practices distilled from audio data case law and international and regional law, as well as ethical principles. While operating ethically is a good in and of itself, the ethicality of CSO conduct can affect the evidentiary value of audio data. (Inter)governmental authorities and judicial institutions, the most likely recipients of audio data for use in accountability processes, may not be allowed to accept audio data that was collected unethically.[11]

2. **The second objective of the Hala Protocol is to align legal requirements and technical specifications.** This is achieved through the technical specifications and resources sections. The digital nature of audio data means that its collection, processing, and transfer takes place in a digital environment. The Protocol aims to help translate between law and tech, giving investigators and legal practitioners the language they need to communicate their requirements to technical experts. To this end, where applicable, some of the Best Practices provide information about tools and software that should or can play a role in operationalising the Best Practice.

3. **The third objective of the Hala Protocol is to help CSOs understand the legal landscape in which audio data collection, processing, and transfer takes place.** This objective is achieved in the Protocol in two ways. On the one hand, Part 2 of the Protocol comprises an extensive Legal Framework section that analyses case law from international criminal courts and tribunals on the treatment of audio data, as well as applicable international and regional laws. On the other hand, each Best Practice contains a condensed legal framework section that highlights the legal principles upon which the practice is based and refers the reader to the corresponding section(s) in the main Legal Framework.

The Scope of the Hala Protocol: What does it (not) cover?

Forms of audio

The Hala Protocol is designed to have a broad scope and to apply to all forms of audio data, regardless of the source of the data and whether the source is open or closed in nature. Audio data is understood as raw or processed electrical signal that is captured and stored in the form of sound, including speech, music, or ambient sound. The Hala Protocol applies to audio data derived from telephone calls, radio communications, voice notes sent over instant messaging apps, audio material posted to social media, and audio material sent over email. The Protocol does not directly address the issue of synthetic audio. While this is an issue of growing concern, the use of AI to generate inauthentic audio content (or other types of audiovisual material) has not yet been widely litigated, meaning that there is insufficient case law to draw on.

[11] See, for example, pages 8 and 9 of the ICC Guidelines Governing the Relations between Court and Intermediaries for the Organs and Units of the Court working with intermediaries (March 2014), which sets out the criteria for selecting intermediaries to work with the Court. A number of the criteria listed are closely related to the ethical principles listed in Best Practice 1 of the Protocol.

The Best Practices reflect the standards articulated by international criminal case law dealing specifically with audio data. That being said, many of the Best Practices are transferable to different types of digital material. CSOs working with other forms of data may therefore find the Best Practices useful across their work.

Only international law

The Protocol does not offer guidance rooted in domestic law, and yet, the collection of some of the above sources of audio may be restricted by domestic law. The Protocol can be used by CSOs to understand the international legal landscape applicable to their work with audio data, but when it comes to the domestic legal landscape, CSOs should seek specialised advice from a domestic legal expert.

Only evidentiary standards

The Hala Protocol is not a 'step-by-step' guide or 'how-to' instruction manual for the collection of audio data.[12] Rather, it is designed to help CSOs align their data practices with evidentiary standards for audio data, increasing the chances of the data being usable as evidence in criminal proceedings.

An example of this limited focus is the Protocol's superficial reference to audio data analysis—understood as engaging with the content of the audio data in order to verify its contents and/or ascertain whether and to what extent it can contribute to proving an element of a crime. The Best Practices that touch upon data analysis are Best Practices 4, 11, and 13, but they do so only in the specific ways that can be directly linked to an aspect of evidentiary value.

While the Hala Protocol identifies evidentiary Best Practices, it does not set all-or-nothing standards. If a CSO cannot, or on occasion does not, conduct their data collection in perfect alignment with the Best Practices, this does not render their data without evidentiary value. There may still be valuable information that can be learned from the data and uses to which it can be put in a legal accountability context.

Methodology: How was it developed?

There are two key elements to the Hala Protocol methodology: the choice of material to base the Best Practices on and the choice of terminology for the Best Practices.

The Hala Protocol Best Practices are grounded in international and regional law and case law, including:

- Case law from international and regional courts, including the International Criminal Court (ICC), International Criminal Tribunal for the Former Yugoslavia (ICTY), International Criminal Tribunal for Rwanda (ICTR), the Special Tribunal for Lebanon (STL), and the European Court of Human Rights (ECtHR);
- The key legal documents of international criminal courts and tribunals, including court statutes, rules of procedure and evidence, and internal guidelines and protocols;
- International and regional human rights treaties, such as the International Covenant on Civil and Political Rights (ICCPR) and the European Convention on Human Rights (ECHR);

[12] By contrast, examples of documents that are closer to a manual in this sense include the OHCHR, Berkeley Protocol on Digital Open Source Investigations: A Practical Guide on the Effective Use of Digital Open Source Information in Investigating Violations of International Criminal, Human Rights and Humanitarian Law (2022) and GLAN/Bellingcat, Methodology for Online Open Source Investigations into Incidents Taking Place in Ukraine since 2022 (2022).

- Legal regimes created by regional organisations, such as the European Union's (EU) General Data Protection Regulation (GDPR); and
- Authoritative guidelines and protocols, such as the OHCHR Berkeley Protocol on Digital Open Source Investigations, the PILPG Handbook on Civil Society Documentation of Serious Human Rights Violations, the WITNESS Video as Evidence Guide, and the United Nations Guiding Principles on Business and Human Rights.

As stated above, the Hala Protocol does not consider the specific requirements and restrictions applicable in domestic legal systems. Domestic rules differ significantly across countries and jurisdictions, which makes it difficult (or in some cases impossible) to formulate broadly applicable standards. International instruments and case law provide a helpful baseline to which CSOs can align their audio data practices thanks to their global character.

The terminology used in the Best Practices denotes the degree to which the practice should be seen as obligatory. The guidelines can be seen as falling into one of three categories: 1) compulsory, 2) recommended, and 3) discretionary.

Compulsory best practices are phrased using **'must'** terminology. For example, 'collectors *must* document each step taken in the collection process'. This terminology will be used when there is supporting case law from international criminal tribunals and/or international and regional human rights courts, and where there is no significant divergence between these cases as to the practice to be followed. 'Must' will also be used where there is international law on the subject–namely treaties–and/or regional law, such as legislation from the EU.

Recommended Best Practices are phrased using **'should'** terminology. For example, 'collectors *should* take steps to safeguard the physical and psychological welfare of personnel'. This terminology will be used when there are authoritative guidelines and documents to support the practice, such as the OHCHR Berkeley Protocol on Digital Open Source Investigations, the PILPG Handbook on Civil Society Documentation of Serious Human Rights Violations, or the UN CEB Principles on Personal Data Protection and Privacy. 'Should' will also be used where there is international case law but notable divergence exists within the case law, and where there are international and regional laws but there is divergence between them.

Discretionary Best Practices are phrased using **'consider'** terminology. For example, 'collectors *should consider* seeking cooperation from the host State if this would further their collection efforts'. This terminology will be used for practices that do not have supportive case law, laws, or authoritative guidelines and protocols, but where conversations with practitioners have flagged the practice as relevant and important to consider. Best Practices based on ethical principles also use 'consider' terminology.

Overview of the Protocol: How is it structured?

The Hala Protocol is organised in two Parts. The Best Practices are contained in Part 1, which is divided into five sections. The first section contains general Best Practices that are applicable throughout a data collection effort. This is followed by three sections that reflect different phases in a data collection effort: audio data collection, audio data processing, and audio data transfer. The final section of Part 1 comprises

Best Practices that are specific to particular audio data collection methods, namely data download, data scraping, and radio interception.

The Protocol is not designed to be read chronologically. Depending on a CSO's workflow, different Best Practices will be relevant at different times, and workflows can be linear, cyclical, or otherwise organised. In this respect it is worth keeping in mind that the Protocol is not a step-by-step guide.

Part 2 of the Protocol contains the Legal Framework that underpins the Best Practices. After providing a high-level overview of the applicable legal frameworks, Part 2 details how international human rights law applies to audio data collection by CSOs. Particular attention is paid to the right to privacy and data protection and the right to a fair trial. This is followed by an in-depth analysis of key evidentiary concepts in international criminal law, and how these concepts—admissibility, relevance, probative value, prejudicial effect, and weight—are to be understood in the context of audio data.

The Hala Protocol concludes with a glossary of terms.

PART 1: BEST PRACTICES

The Best Practices (BPs) are structured as follows:

- **BP Title/Action Statement:** setting out the action to be taken by the Collector;
- **Explanatory Note:** detailing and clarifying the action statement;
- **Technical Specifications + Resources:** list of resources, tools, and techniques to be considered when operationalising the BP;
- **(Condensed) Legal Framework:** indicating each BP's foundation in law with hyperlinks to the relevant areas of the Protocol's Part 2 'Legal Framework', plus the applicable ethical principles.

Words and phrases in **bold** are defined in the glossary. Definitions are not included in the Explanatory Note for the sake of brevity. The authors recommend reading the BPs with the glossary at hand.

General Best Practices

This section includes the Best Practices that apply to the totality of the data **collection effort**, from the pre-planning stage, through to the transfer of data to a third-party recipient such as an **accountability mechanism**. These General Best Practices should be integrated into all stages of the **Collector**'s workflow and revisited regularly to ensure the safety and security of all persons involved and, overall, strengthen the collected **audio data**'s **evidentiary value**.

1. Operate in accordance with ethical principles.

Ethical principles are the foundational benchmarks for evaluating a **collection effort's** activities.

If **audio data** is used as **evidence** in criminal proceedings, the nature of the data as well as the methodology by which it was collected will be scrutinised. Collection efforts that are conducted ethically are more likely to withstand challenges. Moreover, Collectors that are known to operate unethically or without integrity may experience challenges in establishing relationships with **accountability mechanisms**.

Therefore, the collection effort should be undertaken in accordance with the following ethical principles:

a. *Do No Harm*: Collection activities should be designed and carried out in a manner that takes every necessary effort to avoid and refrain from causing harm, including harm towards the integrity of the data collected, **personnel** involved in the collection effort, **data subjects**, local or otherwise affected communities, or any current or future investigation or prosecutorial efforts.

b. *Legal awareness*: The work of collecting audio data does not take place in a legal vacuum. Yet, in a conflict context, it is not uncommon for the applicable local law to be unclear or ambiguous. The Collector should familiarise themselves with all domestic, regional, and international law applicable to their collection effort and assess the risk of incurring liability under these legal frameworks. The Collector should seek advice from a qualified legal professional wherever needed. Any collection activities that are directed from, or in some way touch upon, EU jurisdiction should also seek out specialised GDPR advice.

c. *Accountability*: Transparency in the methods and results of a data collection effort makes these efforts more accountable. This is achieved by maintaining an **audit trail**, such that the steps may be reproduced to achieve the same outcome. Personnel involved in collection efforts should be aware that they may be called upon to make a sworn written or testimonial statement about this process. The nature of the statement will differ depending on the form and substance of the evidence, the accountability mechanism in which the statement is sought, and the role of the personnel involved. For example, an analyst may be asked about how the data was filtered and organised, what knowledge has been derived from the data, and how that analysis was made; a technical expert or investigator may be asked about the tools used in the collection process and the steps taken to secure the data's **chain of custody**.

d. *Competency:* The Collector should possess or have access to the appropriate resources, equipment, and expertise necessary to carry out the collection effort. If a Collector does not have the capacity to effectively and securely collect and store the expected data, it should consider carefully whether or not proceed with the collection effort. All personnel assigned to a task should possess the necessary skills, and/or adequate supervision, for the work they are undertaking.

e. *Accuracy, Impartiality, and Objectivity*: The quality of the collection effort will depend on the extent to which it is demonstrably conducted in an accurate, impartial, and objective way. The collection effort should prioritise truthfulness, and any weaknesses in the data, including in the **metadata**, should be evaluated, documented, and mitigated/ameliorated to the extent possible. The Collector and its personnel should be cognisant of, and mitigate the impact of, the potential for personal, cultural, and structural biases to affect their work, for example in situations where a team consists predominantly or solely of nationals of one of the States involved in a conflict. Personnel should treat both potentially **inculpatory** and **exculpatory** information they have retained with equal care and consideration. Analysis of the audio data should refrain from exaggerated or overstated opinion or speculation about the potential use of the data in investigations or accountability processes.

f. *Consent*: Documented **informed consent** should be obtained from the people whose voices or other **personal data** are included in the audio data ('data subjects'), where it is possible to do so. For example, if the Collector conducts an in-person audio interview with a witness to a crime, the Collector should obtain the witness' written or recorded consent prior to recording the audio. If the audio was recorded by a third-party, the Collector should request the witness' written or recorded consent from the third party.

Informed consent involves providing the data subject with clear and understandable information, including how the data will be used, whether any risk to the data subject could occur as a result of the audio data's future use, any safeguards put in place by the Collector to minimise said risk, and how data subjects can revoke consent once granted. Informed consent may be forward-looking, allowing the data subject to consent to the data's potential future use(s), for example as evidence in an ongoing or future criminal

proceeding. In this respect, the Collector should consider any predictable accountability pathways and obtain the appropriate form of consent for those pathways.

The Collector should evaluate the possibility of obtaining consent on a case-by-case basis and document the evaluation and conclusion reached.

Obtaining informed consent may not be possible where the data subject is unavailable to provide consent or cannot be found. For example, it may not be possible to find or contact the data subjects of collected open-source audio. Obtaining consent may also not be possible where doing so would be detrimental to the collection effort. For example, if the collected audio data is military communication intercepted over a radio frequency, informing the data subject of the interception may pose a security risk to the collection effort.

In circumstances where consent can and has been obtained, prior to transfer of the data the Collector should consider whether it is necessary and feasible to (again) obtain informed consent from the data subject, particularly if there has been a change to the risk posed to the data subject since the consent was obtained. If obtaining consent is possible but it has been refused or withdrawn by the data subject, the refusal or withdrawal should be documented and the Collector should not use or transfer the data.

The above list reflects a minimum standard of ethics and is not intended to serve as an exhaustive list of applicable ethical principles. Additional ethical principles may apply depending on the respective collection context.

Technical Specifications + Resources

The ethical principles are drawn from the following guiding sources:
(i) Global Rights Compliance, Basic Investigative Standards for Documenting International Crimes in Ukraine (2023);
(ii) UN Inter-Agency Standing Committee (IASC), Operational Guidance on Data Responsibility in Humanitarian Action (2023);
(iii) OHCHR/UC Berkeley Human Rights Centre, Berkeley Protocol on Digital Open-source Investigations: A Practical Guide on the Effective Use of Digital Open-source Information in Investigating Violations of International Criminal, Human Rights and Humanitarian Law (2022) ('OHCHR, Berkeley Protocol');
(iv) ICC/Eurojust, Documenting International Crimes and Human Rights Violations for Accountability Purposes: Guidelines for Civil Society (2022);
(v) PILPG, Handbook on Civil Society Documentation of Serious Human Rights Violations: Principles and Best Practices (2016); and
(vi) The Folke Bernadotte Academy and The Swedish National Defence College, A Handbook on Assisting International Criminal Investigations (2011).

Regarding the ways in which unethical conduct may affect a Collector's relationship with accountability mechanisms, see e.g., the ICC Guidelines Governing the Relations between the Court and Intermediaries

(2014), pages 7-9, which lists criteria for counsel or organs of the court to consider when screening a potential intermediary.

Resources on Accountability
Refer to BP 3 *Technical Specifications + Resources*.

Resources on Consent
For a draft Informed Consent form, see e.g., ICC/Eurojust Guidelines, Documenting International Crimes and Human Rights Violations for Accountability Purposes: Guidelines for Civil Society (2022), page 40 (Annex 1).

For a discussion of elements of informed consent in detail, as well as considerations for obtaining consent from open source data generators, see e.g., S. Dubberley and G. Ivens, Outlining a Human-Rights Based Approach to Digital Open Source Investigations: A Guide for Human Rights Organizations and Open Source Researchers (2022), pages 27-28.

For considerations for obtaining consent in the context of conflict-related sexual and gender-based violence, see e.g., Institute for International Criminal Investigations, Global Code of Conduct for Gathering and Using Information About Systematic and Conflict-Related Sexual Violence (2020) ('Murad Code'). For guidance on applying the Murad Code to the context of online digital open-source information, see Open-Source Practitioner's Guide to the Murad Code (2025) (Note: the linked resource is a pilot version and may be subject to change).
Also see e.g., A. Koenig, et al., Merging Responsibilities: Ethical Considerations for Securing Consent in Open-Source Investigations of Conflict-Related Sexual Violence (2024), which provides a methodology to determine 'whether consent is needed and from whom, who should seek that consent, how the consent should be secured, and when', in the context of digital open source investigations that touch upon conflict-related sexual and gender based violence.

For practical steps that humanitarian organisations can implement to process the personal data in their possession in compliance with personal data protection principles and requirements, see e.g., ICRC, Handbook on Data Protection in Humanitarian Action (2024), which identifies five data protection principles: fairness and lawfulness of processing; purpose limitation; proportionality; data minimisation; and, data quality.

Legal Framework

See section 4.2.4.B. for an outline of the GDPR's scope.
See section 4.3. for the legal basis of treating inculpatory and exculpatory information equally.
See section 5.6. regarding the need for personnel to be aware that they may be called to testify about their work.

2. Operate in accordance with the principle of data minimisation.

Data minimisation is a methodological principle limiting the collection and **processing** of data to that which is adequate, relevant, and necessary to achieve the purpose of the **collection effort**. This purpose should be clearly stated in the Collection Plan (see BP 6).

The **Collector** should ensure it is familiar with the applicable domestic legal framework as it pertains to data minimisation. If the collection effort falls under the purview of the GDPR, then the effort must implement data minimisation.[13] Adhering to data minimisation as a best practice, regardless of whether or not it is required by law, increases the chances that any infringement on privacy in the data collection effort will be considered proportionate and therefore justified. For example, certain forms of **audio** collection, such as data **scraping** or collecting radio **signals**, are more likely to be used to collect a broad and voluminous array of data. As a result, they have the potential to include a high volume of data that is not relevant for the purpose of the collection effort (see BP 13 for guidance on conducting a **relevance** assessment). These sources of audio pose a greater risk of collecting and/or processing data that is not limited to the purpose of the collection effort. This risk can be managed by prioritising data minimisation (see BP 5 on privacy; see BP 26 on data minimisation when data scraping).

Data minimisation should occur throughout the collection effort. It may involve, for example, targeting only relevant sources, locations, or timeframes for collection (discussed in BP 6); using either manual or algorithmic filtering to flag for **deletion** any data collected unnecessarily (see BP 16); and, reviewing the remaining collected data to ensure its relevance (see BP 13).

Additional benefits of data minimisation include, first, that it may help when structuring a collection effort dealing with voluminous data, as a data minimisation approach will require the data to be filtered, assessed, and organised; and, second, that it may help to cut down the required storage volume and therefore the cost of the collection effort, as storage costs can be significant when working with a large quantity of data.

The Collector should ensure that all steps taken towards data minimisation do not undermine the equal treatment of both potentially **inculpatory** and **exculpatory** data (see BP 10).

Technical Specifications + Resources

For a discussion of the ethical considerations when relying on algorithmic filtering, see e.g., Data Science and Ethics Group, A Framework for the Ethical Use of Advanced Data Science Methods in the Humanitarian Sector (2020), pages 29-32.

[13] GDPR, Article 5(1)(c).

Legal Framework

See section 4.2.2.C. on the role that the principle of data minimisation plays in protecting the right to privacy. Observing data minimisation increases the chances that a data collection effort's infringement on privacy will be considered proportionate and therefore justified.
See section 4.2.4.B on 'Scope of the GDPR' and 'Data Protection Measures' under the GDPR.
See section 4.3. on the fair trial right of ensuring that the accused is given exculpatory material.

3. Ensure the collection effort is properly documented.

Adequate documentation in the context of a **collection effort** refers to the creation and retention of a verifiable record of all activities undertaken throughout the collection effort. The documentation of the collection effort should serve as an **audit trail**, in that it should be thorough enough to allow for all steps to be tracked and for the final results to be repeated or reproduced.

Thoroughly documenting the methodology and execution of the collection, **processing**, and transfer of **audio data** is paramount for demonstrating the data's **authenticity**, **reliability**, and **chain of custody** if tendered as **evidence**. The documentation of a collection effort is also used in criminal proceedings to scrutinise whether the collection effort was conducted in a manner that was legal, impartial, and otherwise ethical (or, if not, whether this has tainted the evidence and rendered its admission **prejudicial** or otherwise inadmissible). Thus, policies core to the collection effort, such as the Collection Plan, should be documented in writing and securely stored to prevent tampering or **deletion**/destruction (see BP 6).

All **personnel** who are involved in the collection effort must document their activities and observations.[14] All documentation should be dated, and should indicate the author as well as any supervisors who sign-off on the documentation. Whether done digitally or manually, all documentation should be stored in an organised and structured manner (see BP 11).

Personnel should undertake to carry out two forms of documentation: 1) logging all collection, processing, and transfer activity into a data tracking system, whether manually or automated; and, 2) logging informal contemporaneous notes and recollections, sometimes referred to as a 'mission diary'. These informal notes or mission diaries have been regarded by courts as contemporaneous written records of a collection effort and can be used to corroborate the formal logs of activity in the data tracking system. If the documentation cannot be made contemporaneously, then it should be made as soon as possible after the activity. In the event a member of personnel is called to deliver a sworn statement (whether written affidavit and/or oral testimony) about their role in the collection effort, a relevant mission diary can corroborate their statement and serve as an essential part of the evidentiary assessment of a particular audio recording.

A court may scrutinise not only the information that was documented, but also any perceived gap in documentation.

[14] *Prosecutor v Ongwen* (ICC), Trial Judgment, para. 658; *Prosecutor v Ongwen* (ICC), Confirmation of Charges, para. 51; *Prosecutor v Ongwen* (ICC), Transcript, para. 44, lines 8-24; *Prosecutor v Tolimir* (ICTY), Judgment, para. 64, referring to *Prosecutor v Tolimir* (ICTY), Transcript, page 5033; *Prosecutor v Blagojević and Jokić* (ICTY), Decision on Admission of Intercept Materials, para. 21; *Prosecutor v Katanga and Chui* (ICC), Decision on Bar Table Motion, para 30. See the discussion in sections 5.2. and 5.3. of the Legal Framework.

Personnel must at all times strive to document the collection effort in a manner that is as consistent, clear, and transparent as possible.[15] To this end, the **Collector** may consider establishing a standard form of documentation to be used by all personnel.

Technical Specifications + Resources

Documentation models

For examples of documentation formatting (focused on open-source information), see e.g., OHCHR, Berkeley Protocol (2022), Annex IV 'Online Data Collection Form', and GLAN/Bellingcat, Methodology for Online Open Source Investigations into Incidents Taking Place in Ukraine since 2022, Annex VIII 'Incident Assessment Template' and Annex IX 'Uwazi Fields'.

Documentation tools

If using the open-source programming language Python to process audio, tools such as Sphinx or MkDocs can be used to generate documentation from the Python code.

If using the open-source tool Audacity to process the audio, a tool such as Macros can be used to document changes made to the audio, as seen in this example.

Documentation resources

For minimum documentation requirements for chain of custody, see e.g., General Principles of Digital Evidence, a supplement to the Uniform Principles and Guidelines for Investigations from the Conference of International Investigators (2021), page 3.

For information about keeping a mission diary, see e.g., Folke Bernadotte Academy and Swedish National Defence College, A Handbook on Assisting International Criminal Investigations (2011) page 51.

For information on the technical attributes of audio evidence that may be important to document in order to support audio data authenticity, see e.g., Scientific Working Group on Digital Evidence, Best Practices for Digital Audio Authentication (2017).

Paywall resources

For standardised methods for documenting scientific or technical expert opinions, see e.g., ASTM E620-18 Standard Practice for Reporting Opinions of Scientific or Technical Experts.

Tools such as Hunchly can be used to document a Collector's online activities.

Legal Framework

See section 5.2. on the role of logbooks in determining the relevance of potential evidence.
See section 5.3. on the role of proper documentation for establishing the authenticity and reliability of audio data, as components that contribute to the probative value of audio evidence.

[15] *Prosecutor v Ongwen* (ICC), Trial Judgment, para. 658; *Prosecutor v Ongwen* (ICC), Confirmation of Charges, para. 51; *Prosecutor v Ongwen* (ICC), Transcript, para. 44, lines 8-24; *Prosecutor v Tolimir* (ICTY), Judgment, para. 64, referring to *Prosecutor v Tolimir* (ICTY), Transcript, page 5033; *Prosecutor v Blagojević and Jokić* (ICTY), Decision on Admission of Intercept Materials, para. 21; *Prosecutor v Katanga and Chui* (ICC), Decision on Bar Table Motion, para 30. See the discussion in sections 5.2. and 5.3. of the Legal Framework.

See also, among others, ICC case *Prosecutor v Ongwen* (ICC), Confirmation of Charges, para. 51, where the Chamber, in the context of finding the evidence to be reliable, took note of the detailed explanation of the radio communications interception process provided by the Prosecution.

APPLICABLE ETHICAL PRINCIPLES: Accountability; Accuracy, Impartiality, and Objectivity.

4. Use specific and neutrally descriptive terminology.

In a **collection effort**, specific and neutral terminology refers to language that objectively reflects the content and context in which the **audio data** was collected without the imposition of opinion or bias. For example:

- In describing the nature of the sounds captured in the audio data, an example of neutral wording would be 'a loud, short sound, possibly a gunshot', compared to the non-neutral description of 'the sound of a gunshot';
- In describing the alleged crime(s) captured in the audio data, an example of neutral wording would be 'a possible authorisation of a missile strike', compared to the non-neutral 'authorisation of a missile strike';
- In describing the context of the audio data collection, an example of neutral wording would be 'a gathering of approximately 100 people in the street', compared to the non-neutral 'a large political demonstration';
- In describing the actors involved in the alleged crime(s) captured in the audio data, an example of neutral wording would be 'a group of armed individuals', compared to the non-neutral 'the terrorist group'.

Appropriate and neutral terminology also manages expectations regarding the ultimate use of the data. For example, collected data should not be labelled '**evidence**' unless it has been tendered as evidence before an **accountability mechanism** to prove a fact in question. Rather, it should be referred to as 'potential evidence', 'data', or 'information', which accurately reflects the fact that not all collected data will be relied on as evidence.

Appropriate and neutral terminology should be used with consistency and without exception throughout the collection effort and by all **personnel**, e.g., when documenting the collection effort (see BP 3); when **labelling** audio data files or the content within the files (see BP 11); and in both internal communication among personnel and external communication with third parties. The **Collector** should consider establishing a list of appropriate and neutral terminology that is relevant to the context of the collection effort. This list should be clearly defined in the Collection Plan (see BP 6).

Technical Specifications + Resources

For terminology distinctions, see PILPG, Handbook on Civil Society Documentation of Serious Human Rights Violations: Principles and Best Practices (2016), pages 15-18.

For neutral language relevant to open-source investigations, see GLAN/Bellingcat, Methodology for Online Open Source Investigations into Incidents Taking Place in Ukraine since 2022 (2022), Annex VII: Style Guide and Naming Conventions, page 75.

Legal Framework

See section 4.3. on the presumption of innocence as an element of the right to a fair trial.

APPLICABLE ETHICAL PRINCIPLES**:** Accuracy, Impartiality, and Objectivity.

5. Consider privacy by undertaking regular necessity and proportionality assessments.

There is an interference with the right to privacy whenever **personal data** is collected, processed, and/or transferred. For an interference with privacy to be justified, and thus not a violation of the right, the interference must be necessary to achieve a legitimate aim and the measures taken must be proportionate.[16] Throughout the data **collection effort**, **Collectors** must review the necessity and proportionality of their actions in light of the right to privacy on a regular basis.[17] Whether data is collected in violation of privacy rights can be relevant to a court's assessment of the data's **admissibility** as **evidence**.

Collectors should operate under the presumption that the collected data could contain personal data, and that consequently privacy protections apply. Collectors should regularly review necessity and proportionality even if the data they are collecting will not be reviewed by human eyes, for example if the collection effort's objective is only to collect and store data.[18] This applies equally if the data is collected solely for the purposes of **algorithmic training**.

Reviewing the necessity of a collection effort requires the Collector to ask: could the collection effort objectives be achieved by less privacy-intrusive means? For example, by structuring the collection effort to avoid collecting personal data, or by **anonymising** any collected personal data? If the answer is yes, the collection effort must be adjusted accordingly; if the answer is no, the Collector can then proceed to a proportionality assessment.[19]

The proportionality assessment should be guided by the questions listed in section 4.2.2.C. of the Legal Framework. This list is non-exhaustive and may need to be expanded depending on the collection effort's objectives and specific context.

[16] ECHR, Article 8; Convention 108, Article 11.
[17] *P.N. v Germany* (ECtHR), Judgment, para 85; *Catt v United Kingdom* (ECtHR), Judgment, para 119-120; *Big Brother Watch and Others v. United Kingdom* (ECtHR), Judgment, para. 350, 356; *Case of S. and Marper v United Kingdom* (ECtHR), Judgment, para 119.
[18] ECtHR, Guide to the Case-Law of the European Court of Human Rights: Data Protection (2022), para. 8, citing *Kırdök and Others v. Turkey* (ECtHR), Judgment.
[19] ECHR, Article 8(2); Convention 108, Article 11.

Technical Specifications + Resources

Anonymisation of audio data can be achieved through redaction techniques, including, among others, noise addition, speech transformation, or voice conversion (e.g., zero-shot voice conversion is a means of voice conversation using machine learning that requires no or minimal training data).

Due to the likely sensitive nature of the information undergoing anonymisation, it is preferable for any anonymisation tools to be localised and machine-based, rather than cloud-based, in order to minimise any additional risk to security or privacy.

The Collector should consider the existing scepticism concerning the permanence of anonymisation techniques, particularly if there is other data that can be linked to establish personal data. See e.g., NIST, Interagency Report 8387, Digital Evidence Preservation Considerations for Evidence Handlers (2022), page 11.

Legal Framework

See section 4.2. on when data is classified as personal data and 4.2.1. on the applicability of privacy in military communications.
See section 4.2.2. on the three-part test for determining whether an interference with the right to privacy is justified, and in particular section 4.2.2.B. and section 4.2.2.C. on the necessity and proportionality aspects of this test.

APPLICABLE ETHICAL PRINCIPLES: Do No Harm; Legal Awareness.

Audio Data Collection

Audio data collection can be undertaken in a myriad of ways, depending on the data source and collection technology. The following Collection Best Practices are designed to reduce the risks that can arise to the **evidentiary value** of audio data in the collection effort's preparation and execution. These Best Practices should build upon the General Best Practices in the previous section, and should apply regardless of the collection method used. For Best Practices that are tailored to Specific Data Collection Methods, see BPs 24 - 29.

6. Prepare a plan for the collection, processing, and transfer of audio data files.

The Collection Plan serves as the foundational document for the collection effort. It may be tendered as **evidence** in criminal proceedings in support of the collected **audio data**.

The preparation of a Collection Plan involves developing written documentation prior to the initiation of the **collection effort** that outlines in detail the collection effort's objective(s), context, practical steps from the preparation of collection to the data's transfer to **third-party recipients**, and the foreseeable security risks and how they will be managed. It should include short-term and long-term **preservation** strategies for the collected data.

Once the Collection Plan is established, adherence to the Plan should be closely and continuously monitored and documented. The Collection Plan should be considered a 'living document' and adapted as needed. If the Collection Plan is amended or updated during the collection effort, the changes should be made in writing and communicated to all **personnel**.

The Collection Plan should, in substance, comprise policies for a) data collection, b) data **processing**, c) data transfer, and d) risk management. Each of the policies should provide, at a minimum, detailed answers to the following questions:

a. **Data Collection**
    i. What is the **Collector**'s overall objective? If there are multiple objectives, how are they prioritised?
    ii. For each data collection effort, what is the purpose of the data collection?
    iii. What type(s) of data will be collected, in what formats?
    iv. Who is the intended end user of the data collected?
    v. How will the data, and its **embedded metadata,** be collected? What tools are required?
    vi. What is the contextual framing for the collection effort? (This may include a brief history of the region and/or conflict, relevant political and socio-cultural dynamics, persons of interest, linguistic barriers, etc.)
    vii. What are the domestic, regional, and international legal frameworks applicable to the collection effort? (Advice from a qualified legal professional is recommended.)
    viii. What is the intended timeline of the collection effort?
    ix. How will the collection effort adhere to ethical principles?

x. How will the **informed consent** of witnesses, victims, or other data subjects be included in the data be obtained at the point of collection? What circumstances would make informed consent inaccessible/impossible to obtain?
xi. What types of documentation procedures and tools will be used?
xii. Who are the Collection personnel and what are their roles, qualifications, and responsibilities/authorisations? What is the supervisory and reporting structure?
xiii. Will any cooperation or partnerships be needed to achieve the objectives of the collection effort?
xiv. What are the limitations, weaknesses, and risks (see 'c. Risk Management', below) of the collection effort and its collected data? Are any mitigation steps available?

b. Data Processing

i. What is the applicable legal and regulatory data protection framework? What steps will be taken to ensure compliance? (Advice from a qualified legal professional is recommended.)
ii. How will the data be preserved and stored? What types of long and short-term storage facilities are available? What are the costs of the storage, and other resource considerations?
iii. Who will have access to the data? What criteria will determine access?
iv. How long will the Collector retain the data for, and what infrastructure is required in order to do so?
v. How will the integrity of the data be evaluated, and maintained?
vi. Will the data be **enhanced** in any way? If so, how and for what purpose?
vii. Will the data be analysed? If so, how and for what purpose?
viii. How will the **metadata** be verified for accuracy and completeness, and documented?
ix. How will the data and metadata be organised and **labelled**?
x. Will any data be **deleted**? If so, how will the decision to either delete or retain the data be made, and audited?

c. Data Transfer

i. Will the data be shared? With what **third-party recipients**?
ii. How will assurance be obtained that the data will be safeguarded once transferred? What criteria will be used to conduct a **due diligence** assessment of the recipient?
iii. Will the recipients of the data be required to complete non-disclosure agreements?
iv. In what format will the data be shared, and is the Collector's data format **interoperable** with the intended recipient's?
v. Do the applicable legal framework(s) impose any constraints to sharing data? (Advice from a qualified legal professional is recommended.)
vi. Under what conditions and authorizations will the data be shared? What is the internal decision-making process leading up to data transfer?
vii. Will the tools used to share the data put the integrity or quality of the data at risk? If so, how can the risk be mitigated?

viii. Will the informed consent of the **data subjects** be obtained and documented prior to transfer of the data to a third-party recipient? What circumstances would make informed consent inaccessible/impossible to obtain?

d. Risk Management

i. What are the foreseeable risks of the collection effort to the (physical, cyber, or other) security of the data and the collection equipment; the Collector and personnel; the data subjects and/or affiliated individuals or entities; or other persons/environments, at any point in the collection, processing, and/or transfer of the audio data?
ii. What is the Collector's risk tolerance (i.e. with what level of risk can the Collector cope, and under what circumstances?)
iii. What measures will be taken to prevent and to monitor for a security breach?
iv. What is the procedure in the event of a security breach? To recover from a security breach?

The Collection Plan should be reviewed and updated at regular intervals or prior to any significant changes to the collection effort, based on the most up-to-date best practices.

Technical Specifications + Resources

Data collection and processing resources

For guidance on incorporating data responsibility into organisational data management planning, see e.g., IASC, Operational Guidance on Data Responsibility in Humanitarian Action (2023) ('IASC (2023)') Annex B Template: 'Standard Operating Procedure for Data Management Activity'.

For considerations related to evidence sources, physical media and devices, media longevity, imaging digital data, and storage considerations, see e.g., NIST, Interagency Report 8387, Digital Evidence Preservation Considerations for Evidence Handlers (2022).

For guidance on working with data from mobile devices, see e.g., SWGDE Best Practice for Mobile Device Evidence Collection and Preservation, Handling, and Acquisition (2019).

Data sharing resources

For detailed data sharing guidance, see e.g., IASC (2023) Annex B Template: 'Information Sharing Protocol Template (including a Data Sensitivity Classification)'.

For an example data sharing agreement, see e.g., IASC (2023) Annex B Template: 'Data Sharing Agreement Template'.

Risk management resources

For a detailed how-to resource for civil society organisations to establish essential security policies and protocols, see e.g., C. Guerra Merlo, Safe and Documented for Activism Manual (2018).

For a risk management SOP template, see e.g., IASC (2023) Annex B Template: 'Standard Operating Procedure for Data Incident Management'.

For an overview of online and offline security considerations applicable to open source investigations, see e.g., OHCHR, Berkeley Protocol, 'Security' Chapter, pages 33-41.

For risk management checklists, see e.g., International Electrotechnical Commission (IEC), What Security Topics Should Be Covered in Standards and Specifications for 'a checklist for the combination of standards and specifications used in implementations of systems', and the UN Global Pulse Risk, Harms and Benefits Assessment Tool (2019) minimum checklist for international development and humanitarian organisations to understand and minimise potential risk of harm to individual rights when collecting data.

For guidance on the development of a cyber incident response plan, see e.g., U.S. Department of Justice Cyber Security Unit, Best Practices for Victim Response and Reporting of Cyber Incidents, Version 2.0 (2018); see also e.g., NIST SP 800-30 Rev. 1 Guide for Conducting Risk Assessments (2012) and/or the NIST Cybersecurity Framework 2.0 (2024), which offers a taxonomy of high-level cybersecurity outcomes that can be used by any organisation to better understand, assess, prioritise, and communicate its cybersecurity efforts.

Paywalled resources

For internationally recognised principles, framework, and processes for managing risk, see e.g., International Organization for Standardization (ISO), ISO 31000:2018 - Risk Management--Guidelines (2021); and ISO/IEC 27001:2022, information security, cybersecurity, and privacy protection (2022), which also includes requirements for the assessment and treatment of information security risks tailored to an organisation's particular needs.

For a model cyber security threat analysis, see e.g., A. Shostack, Threat Modeling: Designing for Security (2014), which refers to the STRIDE framework (see e.g., 'STRIDE Chart,' Microsoft Security Blog (2007)), identifies common cybersecurity needs—authentication, integrity, non-repudiation, confidentiality, availability, and authorization—and links them to six corresponding cybersecurity threat categories: spoofing, tampering, repudiation, information disclosure, denial of service, and elevation of privilege.

Legal Framework

See section 4.2.2.B. and section 4.2.2.C. on the factors relevant to an assessment of whether an interference with privacy is necessary and proportionate, and therefore justified.
See section 4.2.4.B. on the provisions applicable to EU connected data under the GDPR, in particular those discussed in the 'Legal Bases for Data Processing', 'Notification Requirements', and 'Data Transfers' subsections.

APPLICABLE ETHICAL PRINCIPLES: Legal Awareness; Competency; Accuracy, Impartiality, and Objectivity; Consent.

7. Consider the feasibility and desirability of establishing cooperation with relevant State authorities and/or non-State entities.

The **Collector** should assess the advantages and disadvantages of establishing relationships with the relevant State and non-State entities in light of the Collector's mission objectives, security strategy, country context, and applicable law. Any such assessment should take place in advance of launching the **collection effort** (see BP 6).

Cooperating with State authorities (inclusive of *de facto* State authorities), or non-State entities, such as **Non-State Armed Groups** (NSAGs), as well as private/commercial entities, may be necessary or beneficial to achieving the objectives of the collection effort. For instance, when a collection effort is taking place on the ground in a conflict area, cooperating with the relevant State/NSAG may assist the Collector in acquiring permission or permits to safely enter and move around the territory; to bring in, install, and operate equipment in the territory; and to leave the territory with the equipment and data once the collection effort is complete. Similarly, entering into contracts with private/commercial entities can allow the Collector to secure property leases or other arrangements that may be necessary to the collection effort.

Collectors should also consider that State/NSAG authorities may be hostile to the collection effort, or may themselves perpetrate or contribute to commission of violations in respect of which the data is sought by the Collector. Under such circumstances, cooperation with the host State or with certain State authorities/with the NSAG may not be feasible. The Collector should obtain comprehensive legal advice about the legal landscape within which they seek to operate and the risks involved.

If cooperation is established, the Collector should then assess the feasibility and desirability of formalising the cooperation in writing.

Risks associated with establishing cooperation with States / non-State entities

Establishing links with the authorities of a State or an NSAG involved in an armed conflict may have certain detrimental implications for the Collector. These implications will depend on the mode of cooperation that is established between the Collector and the State/NSAG. For instance, if a Collector cooperates with a State/NSAG in a close manner, the Collector may be regarded as an agent of the State/NSAG. As a consequence, the Collector may become targetable by the State/NSAG's adversary during the armed conflict, potentially posing a threat to Collector **personnel**, the collected data, or other affiliated individuals or entities.

The Collector should also consider that links with a State or an NSAG may give rise to actual or perceived biases in the collection effort, which may undermine the neutrality and credibility of the data collected.

If the Collector shares tactical information or intelligence with a State/NSAG and the information or intelligence is used for commission of crimes by the State/NSAG, the Collector may be implicated in such crimes. Notably, if the Collector shares tactical information or intelligence during an armed conflict to assist the State/NSAG in achieving its military objectives, the Collector may be seen as directly participating in the armed conflict, meaning that the Collector's personnel and property, as civilians and civilian objects, would temporarily lose their protection from targeting under international humanitarian law.

The Collector should similarly consider and document any potential risks that may be associated with establishing cooperation with private/commercial entities that are operating within the context of an armed conflict.

The Collector should additionally consider whether its relationship with a State, NSAG, or other entities involved in or party to an armed conflict has the potential to pose any risk of physical, cyber, or legal harm for third parties to the collection effort. For example, the Collector's partners/donors/affiliates, which may comprise private entities, other States, or international or regional institutions, may wish to avoid being seen to assist one or another party to the respective conflict.

Accordingly, the Collector should carry out the necessary **due diligence** before establishing any cooperation with a State, NSAG, or private/commercial entity and, if necessary, create safeguards to protect against the above risks.

Technical Specifications + Resources

Refer to BP 6, *Technical Specifications + Resources*, 'Risk management resources'.

For information on how to carry out necessary due diligence prior to establishing cooperation with a private/commercial entity, see e.g., OHCHR, UN Guiding Principles on Business and Human Rights (2011), and the steps outlined in OHCHR, UN Guide on Heightened Human Rights Due Diligence for Business in Conflict-Affected Contexts (2022), Annex A 'Heightened Human Rights Due Diligence - Q&A for Businesses' and Annex B 'Heightened Human Rights Due Diligence - action overview'.

Legal Framework

See section 4.1. on the duties applicable to private entities and for actors operating in conflict contexts. See section 3.3. for an overview of the field of international humanitarian law, and section 4.2.1.B. for discussion of how participation in hostilities affects civilian status.

APPLICABLE ETHICAL PRINCIPLES: Do No Harm; Legal Awareness.

8. Assess the collection effort's tools and techniques for possible impact to the integrity of the collected data.

The **Collector** should be familiar with any possible risk posed to the integrity of the data by any equipment or software used and its form of use (hereinafter 'tools and techniques') throughout the **collection effort**. Various tools and techniques may have impairing effects, including by altering or degrading the **audio data**, stripping its **metadata**, or generating inaccurate metadata. From an evidentiary standpoint, these outcomes could undermine the data's **authenticity** and **reliability**.

The tools and the techniques used throughout the collection effort should be validated to ensure they will not impact the data's integrity. The validation of new tools and techniques should take place prior to their

use on collected audio data, and should be done on known **data sets** so that any inconsistencies can be identified. Any possible or known impairment or risk of impairment to the integrity of the data must be documented, along with its possible or known cause and any steps taken to cure the problem.[20]

An error in the collection tool's settings can result in inaccuracies in the metadata that is generated by a particular device (the '**embedded metadata**'). For example, a voice recording made with a device that is configured to the wrong time zone may display the incorrect time in the recording's embedded metadata. The incorrect time of recording embedded in the data could jeopardise the recording's reliability in a criminal proceeding by, for example, casting doubt on when an alleged event occurred. To ensure consistency, the Collector should take steps to ensure the equipment and tools are calibrated to generate accurate embedded metadata. The specifications to which the equipment and tools are calibrated should also be duly documented as part of the **audit trail** and included as a form of **associated metadata** (see BP 17).

External factors may also impair the accuracy of embedded metadata. For example, many devices, such as cell phones, will embed geolocation in metadata. Yet, the accuracy of a phone's metadata could depend on the phone's access to cell service, or be vulnerable to disruptions of GPS technology ('jamming') that are commonplace in conflict zones. If a voice memo is recorded with a cell phone while the device has poor or disrupted GPS connection, the geolocation data attached to the voice memo could be missing or inaccurate—for example, by portraying a different location.

The Collector should consider regularly auditing the quality of the audio data and the accuracy of the embedded metadata.

Technical Specifications + Resources

For a list of points and questions to consider when deciding on new tools, see e.g., OHCHR, Berkeley Protocol, Annex V 'Considerations for validating new tools'.

For transferable guidance on tool and technique validation, see e.g., SWGDE's Model Standard Operation Procedures for Computer Forensics (2012), page 6.

For guidance on leveraging the consistency of the specifications to which collection tools are calibrated as a corroborative feature of the embedded metadata, see e.g., ProofMode, Three Layer Problem: Integrity, Consistency, Synchrony.

Certain smartphone applications offer a 'controlled capture' function, meaning the information is captured with comprehensive metadata and stored in a manner that secures its chain of custody. Examples include the EyeWitness to Atrocities app and the Guardian Project's ProofMode app. These apps do not offer a capture option that is specific to audio, however their video functions may be used to record audio.

[20] This is an extension of the requirement stated in BP3 that Collector 'personnel must at all times strive to document the collection effort in a manner that is as consistent, clear, and transparent as possible': *Prosecutor v Ongwen* (ICC), Trial Judgment, para. 658; *Prosecutor v Ongwen* (ICC), Confirmation of Charges, para. 51; *Prosecutor v Ongwen* (ICC), Transcript, para. 44, lines 8-24; *Prosecutor v Tolimir* (ICTY), Judgment, para. 64, referring to *Prosecutor v Tolimir* (ICTY), Transcript, page 5033; *Prosecutor v Blagojević and Jokić* (ICTY), Decision on Admission of Intercept Materials, para. 21; *Prosecutor v Katanga and Chui* (ICC), Decision on Bar Table Motion, para 30. See the discussion in sections 5.2. and 5.3. of the Legal Framework.

Note: There are no known and publicly available controlled capture applications specific to audio data at the time of this Protocol's completion, although such tools are known to be in development.

Legal Framework

See section 5.2. and section 5.3. on the importance of accurately documenting the date and circumstances of the collection of information when establishing the relevance and probative value of evidence.

See also *Prosecutor v Bemba* (ICC), Public Redacted Version of 'Decision on the Prosecution's Application for Admission of Materials into Evidence Pursuant to Article 64(9) of the Rome Statute' of 6 September 2012, para. 84, wherein the Trial Chamber refers to the role of 'date, circumstances and context in which the recording was created' in establishing relevance and probative value.

APPLICABLE ETHICAL PRINCIPLES: Accountability; Accuracy, Impartiality, & Objectivity.

9. Request the original copy and metadata of any audio data received from a third party source, if practicable.

A **collection effort** may involve audio data received from **third-party sources.** For example, when third-party individuals or organisations on the ground in a conflict context do not have the capability to adequately store, process, or **preserve** the **audio data** they are collecting, they may elect to send the data to a **Collector** with these capabilities.

When Collectors receive audio data from third-party sources, they should consider taking steps to ensure the received audio data is neither **enhanced** nor degraded, whether by a third party or as a result of the transfer of the data. For example, data may be degraded if it has gone through data compression (see BP 18), or it may become degraded if the audio formatting used by the source and the Collector is incompatible. The Collector should aim to identify any such vulnerabilities that would cause the audio data, if put forth as **evidence**, to be considered inadmissible or be granted less **weight** if admitted. If such vulnerabilities exist, the Collector must clearly document them.[21]

Ideally prior to receipt of the audio, the Collector should consider requesting from the third-party source, if available:

1. the **original copy** of audio data,
2. sent via an end-to-end **encrypted** communication platform,
3. inclusive of all available **embedded metadata** and **associated metadata**.

[21] This is an extension of the requirement stated in BP3 that Collector 'personnel must at all times strive to document the collection effort in a manner that is as consistent, clear, and transparent as possible': *Prosecutor v Ongwen* (ICC), Trial Judgment, para. 658; *Prosecutor v Ongwen* (ICC), Confirmation of Charges, para. 51; *Prosecutor v Ongwen* (ICC), Transcript, para. 44, lines 8-24; *Prosecutor v Tolimir* (ICTY), Judgment, para. 64, referring to *Prosecutor v Tolimir* (ICTY), Transcript, page 5033; *Prosecutor v Blagojević and Jokić* (ICTY), Decision on Admission of Intercept Materials, para. 21; *Prosecutor v Katanga and Chui* (ICC), Decision on Bar Table Motion, para 30. See the discussion in sections 5.2. and 5.3. of the Legal Framework.

Embedded metadata

A digital asset may lose its embedded metadata when it is posted online or shared through some messaging services. The Collector should therefore:

1. Request that the original copy of the audio data be sent through a communication channel that does not strip embedded metadata, where possible;
2. Identify what metadata is embedded in the received audio data; and
3. Verify the existing embedded metadata with the third-party source to identify any gaps or inconsistencies (see BP 8).

Associated metadata (discussed in greater detail in BP 17)

The Collector's inquiry to the third-party source should include, but is not limited to, a request for:

1. Information about the collection process, such as who collected the audio, and for what purpose;
2. The **chain of custody** of the audio data, inclusive of verification features such as the data's **cryptographic hash value** or **cryptographic signature**;
3. Any formal or informal notes taken in relation to the audio's collection;
4. Information about any storage, preservation, or enhancement processes applied to the audio data;
5. Relevant information about the context within which the audio was collected; and
6. If the audio includes human voice(s), the answers to the following questions: was the consent of the **data subject(s)** obtained? If not, why not? Is it possible, as well as appropriate from a security and ethical standpoint, to contact the data subject(s) to obtain their consent to transfer the audio?

Technical Specifications + Resources

Techniques to share audio data without undermining its metadata include

- using a sharing service that does not strip an audio file of its embedded metadata during transfer, e.g., ProtonDrive; and/or
- wrapping the audio file(s) in a separate, non-destructive container, e.g., a ZIP or TAR archive, prior to sharing.

When using an instant messenger service, e.g., WhatsApp, sharing an audio file as a document attachment preserves embedded metadata.

For guidance on identifying and analysing an audio file's metadata, see e.g., SWGDE Best Practices for Digital Audio Authentication (2018), Section 4.4. 'Global Analyses'.

For guidance on reviewing audio data submitted by a third-party source, see e.g., SWGDE Best Practices for the Enhancement of Digital Audio (2020), paras. 2.1-2.4.

For guidance on cryptographically binding provenance metadata to a digital asset, see e.g., the Coalition for Content Provenance and Authenticity's C2PA Technical Specification and C2PA Explainer.

Legal Framework

See section 4.2.3. on protecting the right to privacy when audio data was collected by a third party.
See section 5.2. and section 5.3. on the importance of accurately documenting the date and circumstances of the collection of information when establishing the relevance and probative value of evidence.

See also *Prosecutor v Bemba* (ICC), Public Redacted Version of "Decision on the Prosecution's Application for Admission of Materials into Evidence Pursuant to Article 64(9) of the Rome Statute" of 6 September 2012, para. 84, wherein the Trial Chamber refers to the role of 'date, circumstances and context in which the recording was created' in establishing relevance and probative value.

APPLICABLE ETHICAL PRINCIPLES: Legal Awareness; Accountability; Accuracy, Impartiality, and Objectivity.

Audio Data Processing

The term 'data processing' covers a breadth of different activities, including the organisation and storage of data, the **enhancement** of data, the **preservation** of data, and more. As data moves through different **processing** stages, there are a number of points at which its **evidentiary value** may be compromised. The Audio Data Processing Best Practices, below, are designed to help **Collectors** minimise this risk. During the processing of data there are also several opportunities to maximise the data's evidentiary value.

10. Treat potentially exculpatory and inculpatory data equally.

At certain criminal **accountability mechanisms**, such as at the International Criminal Court (ICC), the Prosecutor has a statutory obligation to investigate **inculpatory** and **exculpatory** circumstances equally. The classification of **evidence** as exculpatory or inculpatory will depend on a number of factors, including the charges confirmed and the identity of the potential suspects and accused.

Looking ahead to collected **audio data**'s potential use as evidence, the **Collector** should ensure audio data is treated neutrally once it is collected and retained, i.e. all data should be treated in the same manner, regardless of whether it is considered as potentially inculpatory or exculpatory. 'Treat' in this context applies to the way in which the data is handled after collection, specifically in the processing phase of the **collection effort**. The Collector should bear in mind any potential suspect or accused's fair trial guarantees, as well as issues regarding **equality of arms**, and should implement safeguards against the existence or perception of bias in the collection effort.

The Collector should ensure its **personnel** are adequately trained to treat all data equally. The Collector should moreover ensure that any data that is marked for **deletion** is thoroughly reviewed, in accordance with BP 17. This is designed in part to prevent potentially inculpatory or exculpatory audio data from being deleted.

Legal Framework

See section 4.3. detailing the rights of accused persons in connection with the right to a fair trial, in particular the right to be given exculpatory material.
See also Article 54(1)(a) of the ICC Statute, which provides the duties of the ICC Prosecutor to investigate incriminating (inculpatory) and exonerating (exculpatory) circumstances equally: 'The Prosecutor shall (...i)n order to establish the truth, extend the investigation to cover all facts and evidence relevant to an assessment of whether there is criminal responsibility under this Statute, and, in doing so, investigate incriminating and exonerating circumstances equally.'

APPLICABLE ETHICAL PRINCIPLES: Accuracy, Impartiality, and Objectivity.

11. Organise audio data to ensure it is findable, verifiable, and reviewable.

Collectors should employ an organisational asset management system for the **audio data** that is suited to the type and volume of data and to the objective of the **collection effort**. The organisational system should meet three needs: the data must be 1) searchable and findable; 2) unique and verifiable; and, 3) reviewable.

First, to ensure data is searchable and findable, data should be effectively **labelled**. For example, data could be labelled with the time, date, and location of collection or with a unique alphanumeric identification number. Analysts can also assign one or more labels to an audio recording to signal the relevant aspects of the audio's content. For example, a recording that appears to discuss troop movements could be labelled with 'possible troop movements'. This allows for audio data to be filtered and thus findable based on its content or by the values of its metadata.

The labelling scheme used by the Collector should be appropriate for the type and volume of data collected. The scheme should be clearly documented, communicated to all **personnel,** and applied consistently, with training provided if necessary. The Collector should avoid using a labelling scheme where the nomenclature used reaches a maximum value and prohibits further labelling within the same framework. The Collector should also consider whether the labelling scheme is comprehensible to, and **interoperable** with, foreseeable **third-party recipients** of the data. The labelling should be included as part of the audio data file's **associated metadata**.

The searchability and findability of audio data is further ensured by linking data together. The organisational system should provide for the creation of **audio data files**, which group together the **original copy** of the audio data, **duplicates** (if made), and either include or link to the **metadata** (see BP 17), such as the records documenting the collection effort (see BP 3). If different audio data files contain mutually corroborating information—for example, if two collection devices recorded the audio at the same time, date, and place—these should be linked within the organisational system. Creating clear linkages between audio data and associated or corroborative data is important for assessing the **reliability** and **authenticity** of the data, and thus potentially maximising the data's **probative value** in the event that it is presented as **evidence** in court.

Second, a well-designed organisational system will help ensure audio data is unique and verifiable. Uniqueness is key to preventing accidental overwrites or duplication; verifiability permits demonstration of the data's **chain of custody** from the point of collection onward. To achieve this, **preservation** information generated in accordance with BP 14, such as **cryptographic hash values** and **cryptographic signatures**, should be clearly linked with the audio data.

Third, the organisational system must facilitate regular reviews of the data held by the Collector in order to determine whether it is still necessary to retain the data (see BP 2), and whether its retention is in line with privacy protections (see BP 5).[22] While it may not be feasible to know the content of every audio recording collected, Collectors should at a minimum have an overview of the data they have collected and stored.

[22] ECtHR, Guide to the Case-Law of the European Court of Human Rights: Data Protection (2022), para. 205, identifies cases that support the need for regular review of data retention periods and practices, including *Gardel v France* (ECtHR), Judgment, para. 69 and *Peruzzo and Martens v Germany* (ECtHR), Decision, paras. 44-49.

Technical Specifications + Resources

For further information on establishing a system that will allow for the verification of the audio data's integrity and provenance, see e.g., the Coalition for Content Provenance and Authenticity's C2PA Explainer.

For guidance on the organisation of metadata, drawing from digital audio archiving, see e.g., IASA Technical Committee, Guidelines on the Production and Preservation of Digital Audio Objects (2009), Section 3: 'Metadata', as well as e.g., ISO/TC 46/SC11N800R1, Building a metadata schema – where to start.

For guidance on standards-based hashing algorithms and a description of which hashing algorithms are secure, see e.g., Federal Information Processing Standards Publication 'FIPS PUB 180-4', as well as e.g., SWGDE Position on the Use of MD5 and SHA1 Hash Algorithms in Digital and Multimedia Forensics (2019).

For guidance on hashing generally, such as what steps a Collector should take when a hash comparison fails, see e.g., NIST IR 8387, Digital Evidence Preservation Considerations for Evidence Handlers (2022), page 7.

The choice of hashing algorithm should be periodically reassessed and revised as needed to take advantage of applicable advances in cryptography.

A number of tools and approaches exist for the production of digital signatures, including the maintenance of a Public Key Infrastructure (PKI), which uses trusted Certificate Authorities (CAs) to issue certificates authenticating identities and securing communications. For more information on PKIs and CAs see e.g., the United States government IDManagement.gov Explainer on the topic.

An alternative to the centralised authorities found in PKI structures are Decentralised Identifiers (DIDs), which are self-sovereign, independent, cryptographically verifiable identities. For more about this novel technology, see e.g., C. Mazzocca et al, A Survey on Decentralized Identifiers and Verifiable Credentials, Arxiv (2024); also see e.g., the World Wide Web Consortium (W3C) Verifiable Credentials Use Cases and their Specification.

Legal Framework

See section 4.3. on the presumption of innocence as an element of the right to a fair trial.
See section 5.2. and section 5.3. on the importance of corroborative information when establishing the relevance and probative value of potential evidence.
See section 5.3. on the role played by chain of custody in the probative value of potential evidence.

APPLICABLE ETHICAL PRINCIPLES: Accountability; Accuracy, Impartiality, and Objectivity.

12. Consider transcribing and, if necessary, translating any audio data that contains voice.

The **Collector** should consider transcribing collected **audio data** that includes a human voice in order to create a written record that faithfully reflects the content of the audio data. Collected audio data that

includes a human or generated voice may also need to be translated, e.g., when it includes a voice speaking in a language other than that spoken by the **personnel** who will assess the audio data's **relevance**. The Collector may also consider translating the **audio** into the official language(s) of the intended recipient(s) of the audio data, e.g., certain domestic or international courts or tribunals. The Collector should consider taking steps to ensure the translation is thorough, accurate, and impartial.

Transcriptions and translations form a part of the collected audio's **associated metadata** (see BP 17). They should therefore be appropriately included in the relevant **audio data file** (see BP 11). Notably, if the audio has been **anonymised** or partially **deleted**—for example, for privacy reasons per BP 5—the same should be done to the audio's transcription and/or translation.

Both transcription and translation may be conducted automatically by appropriate software or manually by personnel. The tools and techniques used should be thoroughly documented. Per BP 8, the Collector should assess the likelihood of whether its transcription or translation tools or techniques may generate incomplete and/or potentially inaccurate data, or whether using the tools poses any risk to the security of the **collection effort** or the privacy or security of the **data subject**. As part of the collection effort's risk management approach, particular care should be paid to the tool's potential use of cloud-based resources and the associated risks (see BP 6).

If the voice-containing audio is not adequately or wholly intelligible, the resulting transcription and translation may indicate the sections which are unintelligible by, for example, noting /*unintelligible*/ in the text. In the event of possible or known vulnerabilities, such as an inaccurate transcription or translation, the vulnerabilities must be documented along with any measures taken to ameliorate them.[23]

Technical Specifications + Resources

Certain translation and transcription tools (e.g., Whisper, Otter.ai) can automatically, preliminarily estimate whether the audio data is likely to contain a human voice or not. They may also be used to detect the primary language spoken, transcribe the audio in its original language, and/or translate the audio.

The tool(s) used should support the range of languages likely to be spoken in the audio data collected, as well as the language(s) to which audio would have to be translated.

Note: The use of such a translation and/or transcription tool may involve the transfer of the audio outside of the Collector's possession and control, which may pose a risk to the privacy and/or security of the collection effort.

[23] This is an extension of the requirement stated in BP3 that Collector 'personnel must at all times strive to document the collection effort in a manner that is as consistent, clear, and transparent as possible': *Prosecutor v Ongwen* (ICC), Trial Judgment, para. 658; *Prosecutor v Ongwen* (ICC), Confirmation of Charges, para. 51; *Prosecutor v Ongwen* (ICC), Transcript, para. 44, lines 8-24; *Prosecutor v Tolimir* (ICTY), Judgment, para. 64, referring to *Prosecutor v Tolimir* (ICTY), Transcript, page 5033; *Prosecutor v Blagojević and Jokić* (ICTY), Decision on Admission of Intercept Materials, para. 21; *Prosecutor v Katanga and Chui* (ICC), Decision on Bar Table Motion, para 30. See the discussion in sections 5.2. and 5.3. of the Legal Framework.

Legal Framework

See section 5.2. on the importance of translations and transcriptions when establishing the relevance of potential evidence.

APPLICABLE ETHICAL PRINCIPLES: Accuracy, Impartiality, and Objectivity.

13. Assess the relevance of the audio data.

An assessment of **relevance** of **audio data** will likely involve **critical listening** to determine whether the collected **audio** contributes to the objective of the **collection effort**.

For example, if the objective of the collection effort is to collect **evidence** of criminal acts perpetrated during an armed conflict, then information that could indicate the involvement or contribution of actors alleged to be involved in this conduct would be relevant. Information about the circumstances and context in which such acts were committed would likewise be relevant.

The assessment of relevance aids in accurately labelling the audio data and linking it with associated data (see BP 11). If the audio data contains inaudible or unintelligible speech or sounds, it may be necessary to first **enhance** the audio on a **duplicate copy** in accordance with BP 15 so that it can be audibly assessed.

Potentially relevant information uncovered through audio data in the context of an armed conflict may include, *inter alia*:

- The movement and location of armed forces in an area where crimes are committed;
- The factual circumstances of criminal acts that transpired (e.g., how and when the acts took place; a type of weapon, aircraft, or vehicle used and its manner of use; identifying information about key victims, witnesses, or perpetrators, including speech recognition data);
- The command structure and hierarchy of armed forces (which may be relevant to questions of the responsibility of commanders and other superiors over those who are under their effective command and control);
- Any orders issued for the commission of alleged criminal acts;
- Any acknowledgement that an alleged criminal act has taken place;
- Statements that assist in revealing an alleged perpetrator's intent in relation to the commission of a particular crime;
- The identities of military leaders (i.e. squad, platoon, company and battalion commanders) as well as political leaders involved in the alleged commission of a crime; and
- Indications of planning, and discussions or negotiations regarding the alleged perpetration of a crime.[24]

Audio data that is deemed relevant should be duly **preserved** in accordance with BP 14. Audio data that is deemed irrelevant should be marked for **deletion** in accordance with BP 16 and the safeguards outlined therein. Irrelevant audio data includes any audio that can be wholly characterised as 'noise', 'static', or

[24] *Prosecutor v Ongwen* (ICC), Trial Judgment, paras 1618–1638, 1857, 1871, 2001–2005.

'silence' even after enhancement, as well as any audio that does not contribute to the objective of a collection effort. As an additional safeguard, wherever possible, **Collectors** should consider having at least two qualified **personnel** make a documented determination of irrelevance before audio data can be tagged for deletion.

Relevance assessments may not be a practicable step for every Collector, for every collection effort, or for all collected data. However, if the Collector does have the capability or mandate to assess the audio for relevance, then such an assessment should take place as soon after collection as is practicable. In the interim period between its collection and its assessment for relevance, the audio data should be securely stored.

Note: The data's relevance to the collection effort, determined by the Collector, is distinct from the data's relevance to a legal proceeding, which is determined by a court. If the audio data is intended to be used for future civil or criminal proceedings, whether domestic or international, it will only be admitted as evidence if it is deemed relevant to the specific issues in question before the court.

Technical Specifications + Resources

Regarding the need to create a copy or duplicate of the audio prior to its analysis, see e.g., IST Interagency Report 8387: Digital Evidence Preservation Considerations for Evidence Handlers (2022).

For an explanation of the potential relevance of various factual inquiries in a conflict context, see e.g., GLAN/Bellingcat, Methodology for Online Open Source Investigations into Incidents Taking Place in Ukraine since 2022 (2022), page 69, Annex VI 'Factual inquiries and their relationship to the elements of crimes'.

Legal Framework

See section 4.2.2.C. on the importance of data minimisation for ensuring that a collection effort's interference with the right to privacy is proportionate and therefore justified.
See section 4.3. detailing the rights of accused persons in connection with the right to a fair trial, in particular the right to be given exculpatory material.
See section 5.2. on how the term 'relevance' is defined in a legal evidence context.

14. Preserve the relevant audio data and its metadata.

Digital material can be easily **deleted**, lost, corrupted, or tampered with. A well designed **preservation** protocol is therefore crucial for ensuring that **audio data** can be transferred in the future to **third-party recipients**, such as **accountability mechanisms,** and that those recipients can trust the data's **authenticity**. Poor preservation may negatively affect the data's **probative value** in the event that it is presented as **evidence** in court.

In designing a preservation protocol, **Collectors** should account for 1) the preservation of the material, and 2) the demonstration of the quality of preservation.

First, regarding the preservation of the material itself, audio data should be stored in multiple copies on secure servers with access controls to minimise the security risks posed to the data (see BP 6). The use of secure servers can be costly, so Collectors with limited resources should consider forming partnerships with appropriate organisations if they are collecting sensitive data with high security needs, or if the data will need to be stored for an extended period of time. When choosing a third-party server or arranging their own, Collectors should consider the advantages and disadvantages of centralised versus decentralised storage and choose an approach that best corresponds to their mission objectives.

Second, to demonstrate the quality of the preservation, Collectors should utilise **cryptography** to demonstrate audio data's **authenticity**. Cryptographic tools such as **hash values** can be used to show that data is unchanged from when it was collected, making it more **reliable** and improving its potential probative value. Hash values are a part of the audio data's **metadata** and should be included in or linked to the **audio data file** (see BP 12). The other items in the audio data file should also be assigned a hash value, including transcriptions, translations, log books, and any corroborating information; these values can be listed in a hash list. Once the audio data file is compiled, a hash value should also be generated for, and clearly linked with, the audio data file itself.

Other cryptographic tools Collectors can consider using include a **cryptographic signature**, which can serve to both safeguard and demonstrate a digital asset's provenance.

Collectors must ensure that their preservation protocol saves an **original copy** of the audio data in a way that is unchanged from the point of collection,[25] and only permits changes to be made to **duplicates** of the audio data (e.g., including **anonymisation**, per BP 5, or **enhancements**, per BP 15).

The Collector must be able to report on how each piece of audio data was preserved.[26] Such reports should be linked to the audio data in the data organisational system (see BP 11). Such reports can be composed manually or through the use of a tool that generates the reports automatically.

[25] *Prosecutor v Ongwen* (ICC), Trial Judgment, paras. 654-655; *Prosecutor v. Mladić* (ICTY), Decision on Prosecution Motion for Admission of Documents from the Bar Table, paras 11-12.

[26] This is an extension of the requirement stated in BP3 that Collector 'personnel must at all times strive to document the collection effort in a manner that is as consistent, clear, and transparent as possible': *Prosecutor v Ongwen* (ICC), Trial Judgment, para. 658; *Prosecutor v Ongwen* (ICC), Confirmation of Charges, para. 51; *Prosecutor v Ongwen* (ICC), Transcript, para. 44, lines 8-24; *Prosecutor v Tolimir* (ICTY), Judgment, para. 64, referring to *Prosecutor v Tolimir* (ICTY), Transcript, page 5033; *Prosecutor v Blagojević and Jokić* (ICTY), Decision on Admission of Intercept Materials, para. 21; *Prosecutor v Katanga and Chui* (ICC), Decision on Bar Table Motion, para 30. See the discussion in sections 5.2. and 5.3. of the Legal Framework.

Technical Specifications + Resources

The creation of multiple copies should adhere to the '3-2-1 backup rule', which maintains that a redundant backup of digital information includes at least three copies, stored on at least two different media types, with at least one of them stored off-site. For guidance, see e.g., U.S. Cybersecurity and Infrastructure Security Agency (CISA), US-CERT recommendations and Backup Options.

Commercial storage systems (potentially paywalled) include, e.g.,

- Amazon Web Services; Google Cloud; and Microsoft Azure, which are major storage providers;
- Tresorit, Nord, Egnyte, Proton, for systems with advanced security features (such as encryption or zero-knowledge designs);
- Filecoin, Arweave, Storj, PiKNiK, for systems with long-term guarantees of distributed cold-storage availability and integrity (via resilient decentralised architectures);
- or others, including e.g., Oracle Cloud, OVH, Alibaba Cloud, Tencent Cloud, Backblaze, IBM.

For overarching preservation considerations, see e.g., OHCHR, Berkeley Protocol, pages 60 - 62; and NIST IR 8387, Digital Evidence Preservation Considerations for Evidence Handlers (2022), pages 7 - 10.

For audio file formats that are recommended as having the best chances for survival and continued accessibility, see e.g., Library of Congress, Recommended Formats Statement 2024-2025.

Immutable storage products can pose risks as well as advantages in terms of security, protection from accidental deletion/overwrite, and compliance. They may be a feature of a system's commercial offering (e.g., Amazon's Object Lock), or part of a system's cornerstone design (e.g., Filecoin).

Systems of record can be used to publicly disclose select elements of non-critical, non-identifiable metadata (e.g., cryptographic hash values) for the purpose of registering them with a publicly-verifiable timestamp.

Legal Framework

See section 5.3. on the role played by chain of custody in the probative value of potential evidence.
See also *Prosecutor v. Mladić* (ICTY), Decision on Prosecution Motion for Admission of Documents from the Bar Table, paras. 11-12 wherein the Chamber noted that the tendering by the Prosecution of the original audio recordings contributed to finding these to be relevant and probative.
See also *Prosecutor v Ongwen* (ICC), Trial Judgment, paras. 654-655, wherein the Chamber noted that the Prosecutor had provided an original audio recording alongside the enhanced version. This allowed for comparison between the two and contributed to the evidence being found to be reliable.

15. Only make enhancements to a duplicate copy after securely preserving the original copy.

Enhancements must only be made to **duplicates** of the collected **audio data** and not to the **original copy**.[27] Enhancements in this context include, *inter alia*, any anonymisation of the data, per BP 5. The **preservation** of the original copy is imperative to the audio data's **probative value** as potential **evidence**.

[27] *Prosecutor v Ongwen* (ICC), Trial Judgment, paras. 654-655.

It is important to maintain an unprocessed original copy against which duplicates may be compared. Any duplicate copy must at first be identical to the original copy, such that it would share the same **cryptographic hash value** as the original copy, until the duplicate is enhanced. Any duplicate should be labelled as a duplicate and should be linked back to the original copy (BP 11).

If there are observable, material differences between the contents of the original copy and the enhanced copy, this may affect the **admissibility** of the enhanced copy in the event it is tendered as evidence. Consequently, the audio data could fail to be admitted into evidence, or be granted less **weight** at the evaluation stage of an adjudicative proceeding.

If the audio data has been obtained from a third party and it is not established to be an original copy, the **Collector** should take steps to obtain an original copy and its efforts to do so should be duly documented (see BP 9).

Technical Specifications + Resources

Non-material enhancements of audio data may include:

- cleaning audio from excess noise/static (i.e. 'denoising' via compression or echo cancellation);
- normalising the volume level of the audio;
- isolating or highlighting the human voices or other desired sounds (for example, via frequency equalisation).

See e.g., H. Fayyad-Kazan et al, 'Verifying the Audio Evidence to Assist Forensic Investigation' (2021) pages 29 - 30.

On the desirability to obtain the original copy of the audio data, see e.g., SWGDE Best Practices for the Enhancement of Digital Audio (2020), para. 2.4.

For steps to take when a hash comparison between the original and a duplicate fails, see e.g., NIST IR 8387 (2022), page 7.

Legal Framework

See section 5.3. on the need to work with duplicates when enhancing audio data in order to bolster the reliability, and therefore probative value, of this data.
See also *Prosecutor v. Mladić* (ICTY), Decision on Prosecution Motion for Admission of Documents from the Bar Table, paras. 11-12, wherein the Chamber noted that tendering original audio recordings alongside enhanced versions contributed to finding the recordings to be relevant and probative.
See section 5.3. on the importance of detailed record keeping when enhancing audio data for establishing the reliability, and therefore probative value, of this data.

APPLICABLE ETHICAL PRINCIPLES: Competency; Accountability.

16. Delete any audio data that is deemed irrelevant.

Deletion of **audio data**, if not carefully managed, could result in the permanent loss of potential **evidence**. The consequence of losing potential evidence, whether **inculpatory** and **exculpatory**, is that it could jeopardise potential pathways to accountability and lead to accusations from the opposing party of destroying material (see BP 10). The **Collector** should develop a deletion policy that clearly indicates the circumstances under which collected audio data is to be deleted, and the process for doing so.

The Collector must balance its caution to refrain from deleting potential evidence with its obligations regarding data minimisation.[28] Any audio data that is deemed irrelevant to the objective of the **collection effort** should be deleted in order to adhere to the principle of data minimisation (see BP 2). Deleting irrelevant audio data may also be required by applicable privacy law, for example if the audio data includes a human voice or other forms of **personal data**.

If the audio data is deemed irrelevant , following the assessment of **relevance** detailed in BP 13, the data and its respective **audio data file** should be **labelled** as being ready for deletion. This data should be kept separate from the relevant audio data and should be deleted on a regular basis.

The Collector should consider regularly auditing the data that has been marked for deletion in order to ensure that the assessment of irrelevance has been appropriately carried out. The audit serves as a safeguard against accidental or inappropriate deletion of relevant audio data that might serve as potential evidence.

All deletion data, including a record of deletions and all undertaken audits, must be logged in a tracking system.[29] The Collector should consider logging a general description of the deleted material along with its date range and other relevant information.

Legal Framework

See section 4.2.2.C. on the importance of data minimisation for ensuring that a collection effort's interference with the right to privacy is proportionate and therefore justified.
See section 4.3. detailing the rights of accused persons in connection with the right to a fair trial, in particular the right to be given exculpatory material.

17. Add associated metadata to all relevant audio data files.

As with **embedded metadata**, **associated metadata** can include a range of information that serves to establish the **audio data**'s potential **relevance** and **probative value**. The **Collector** should ensure each

[28] Per GDPR, Article 5(1)(c), collection efforts that fall under the purview of the GDPR must implement data minimisation.
[29] This is an extension of the requirement stated in BP3 that Collector 'personnel must at all times strive to document the collection effort in a manner that is as consistent, clear, and transparent as possible': *Prosecutor v Ongwen* (ICC), Trial Judgment, para. 658; *Prosecutor v Ongwen* (ICC), Confirmation of Charges, para. 51; *Prosecutor v Ongwen* (ICC), Transcript, para. 44, lines 8-24; *Prosecutor v Tolimir* (ICTY), Judgment, para. 64, referring to *Prosecutor v Tolimir* (ICTY), Transcript, page 5033; *Prosecutor v Blagojević and Jokić* (ICTY), Decision on Admission of Intercept Materials, para. 21; *Prosecutor v Katanga and Chui* (ICC), Decision on Bar Table Motion, para 30. See the discussion in sections 5.2. and 5.3. of the Legal Framework.

audio data file includes comprehensive associated metadata, which will enhance the **evidentiary value** of the respective audio data in the event it is tendered as **evidence**.

Associated metadata includes, but is not limited to:

- A written description of the content of the audio data, including the circumstances and context of its creation;
- An explanation of the technical processes and equipment involved in the collection of the audio data and any concerns regarding its quality;
- The personal observations of the **personnel** involved in the collection of the data;
- Information about the audio data's history and **chain of custody**, including:
    - dates and names of personnel involved in the data's creation, **processing**, and subsequent access, alongside details of location, activity undertaken, and purpose of the activity;
    - the status of any **duplicates** made, and a clear link to them;
    - a description of any **enhancements** made to the data;
    - applicable **cryptographic hash values** and/or **signatures;**
- If the **audio** includes a human voice, the transcriptions and, if applicable, translations of the audio (see BP 12);
- The data's **labelling** schema and labels (see BP 11).

Technical Specifications + Resources

Refer to BP 11 *Technical Specifications + Resources*, regarding the organisation of associated metadata in the asset management system.

Legal Framework

See section 5.2. and section 5.3. on the importance of accurately documenting the date and circumstances of the collection of information for the relevance and probative value of evidence.
See section 5.3. on the role that metadata can play in establishing the reliability and authenticity, and therefore probative value, of audio data.
See also *Prosecutor v Bemba* (ICC), Public Redacted Version of "Decision on the Prosecution's Application for Admission of Materials into Evidence Pursuant to Article 64(9) of the Rome Statute" of 6 September 2012, para 84, wherein the Trial Chamber refers to the role of 'date, circumstances and context in which the recording was created' in establishing relevance and probative value'.
See also, among others, *Prosecutor v Ongwen* (ICC), Confirmation of Charges, para 51, wherein the Chamber indicated that the detailed explanation of the process of interception and analysis of radio communications provided by the Prosecution contributed to their finding that the evidence was reliable.

APPLICABLE ETHICAL PRINCIPLES: Accountability; Accuracy, Impartiality, and Objectivity.

18. Safeguard the integrity of any relevant audio data that undergoes file size compression.

It may be necessary for the **Collector** to compress **audio data** in order to reduce its file size and facilitate its **preservation** on a server, and/or its digital transfer to **third-party recipients** over the internet or across a virtual network. Yet, file size compression can lower the quality of the **audio data file**–such that pitch, tones, or other audio details cannot be heard. The compression of an audio data file must therefore be undertaken in a manner that ensures the audio is not fundamentally altered and remains **reliable** and **probative** in the event it is presented as **evidence** in court.[30] While the Collector should strive to maintain the audio data in its non-compressed form, this obligation is to be balanced with the practicality of the volume of data collected.

If the quality of the audio data is affected by the compression to such an extent that its content becomes less intelligible, it may be considered less reliable and probative.

Technical Specifications + Resources

The two types of audio compression are 'lossy' and 'lossless':

- with lossy compression, data is lost and cannot be retrieved in its original form;
- with lossless compression, no data is lost but the compressed file uses fewer bits to represent the information. When the file is reopened, the original data is then reconstructed. The displayed image is identical to the original source image.

For a discussion of different compressed audio file formats, see e.g., HigherHZ, Lossless vs. lossy audio: FLAC, WAV, MP3, and other formats, (2021), which notes that 'MP3 or MPEG Audio Layer III is the most popular lossy audio format and still one of the most used overall'.

For an overview of video/audio compression, see e.g., RGB Spectrum, Digital Video and Compression.

For a description of the potential disadvantages of compression and resulting distortion for evidentiary value, see e.g., A. Koenig and Freeman, Cutting-Edge Evidence: Strengths and Weaknesses of New Digital Investigation Methods in Litigation (2022), pages 1246-1248.

Note: There are no known publicly available tools that permit for the transfer of audio without compression at the time of this Protocol's completion, although there is at least one such tool known to be in development.

Legal Framework

See section 5.3. on establishing reliability, and therefore the probative value, of audio data. The legal framework does not discuss compression specifically, but an analogy can be drawn to case law dealing with enhancement of audio evidence and particularly the requirement that enhanced data should not be materially different from non-enhanced copies.

APPLICABLE ETHICAL PRINCIPLES: Do No Harm; Accuracy, Impartiality, and Objectivity.

[30] Analogical application of *Prosecutor v Ongwen* (ICC), Trial Judgment, paras. 654-655.

Audio Data Transfer

The Audio Data Transfer Best Practices, below, pertain to the transfer of the compiled **audio data files** to external partners, namely national law enforcement, international organisations, international courts and tribunals, and other **accountability mechanisms**. These Best Practices may additionally be applicable to, but are not necessarily intended for, the transfer of data to other **third-party recipients**, such as media organisations, private entities, or non-governmental organisations. Following these Best Practices throughout the transfer of audio data files will help to safeguard the **evidentiary value** of the **audio data**, the safety and security of involved **personnel** and other individuals, and respect for and compliance with human rights law.

19. Use encrypted communication channels with third parties.

Any communication between the **Collector** and third parties where sensitive information may be discussed or transferred should take place through **encrypted** channels. Using secure and encrypted messaging channels helps to ensure that unauthorised third parties cannot monitor or interfere with the Collector's communication with authorised third parties. Use of such channels prevents leakages of the collected audio data or of any identifying information of **personnel** or the individuals with whom the Collector is in contact.

Technical Specifications + Resources

Examples of encrypted communication tools are Signal, Vitru, Wire, or ProtonMail.

The Collector should consider the purpose of the encrypted tool, for example whether it is for communication alone or also for the transfer of data. It may be the case that an encrypted messaging tool is suitable for communication, but if used for data transfer would strip the metadata or otherwise undermine the data's integrity (e.g., via compression).

Legal Framework

See section 4.2.4.B on 'Data Protection Measures' under the GDPR.

APPLICABLE ETHICAL PRINCIPLES: Do No Harm.

20. Conduct a curated risk assessment before transferring audio data to a third party.

A risk assessment for data transfer involves the **Collector** conducting **due diligence** on a specific **third-party recipient** to whom the Collector plans to transfer the **audio data**. The Collector should assess whether the recipient will continue to safeguard the audio data, including the audio data's **evidentiary value**.

The risk assessment should additionally identify:

- whether transfer to the third-party recipient is in line with the objectives of the **collection effort**;

- any privacy and security concerns posed to the audio **data subjects** in the course of or as a result of the transfer; and,
- any risk posed to the integrity of the audio data in the course of or as a result of the transfer.

If the GDPR applies to the collection effort, the Collector must be sure that equivalent privacy protections will be guaranteed by the third-party recipient.[31]

The Collector should carry out a risk assessment prior to transferring audio data to each third-party recipient. A risk assessment should also be carried out in the event of a material change in an existing third-party recipient's capacity or circumstances, such that it could affect that party's ability to safeguard the data and its evidentiary value.

Any such assessment should be documented in writing per BP 3 and **preserved** by the Collector. The assessment should include a clear delineation of what information can be transferred safely to the third party. If the risk assessment establishes that transferring the data would pose a risk to the data, associated individuals, or other aspects of the collection effort, then the data must not be transferred unless the risk can be managed[32] (for example, by **redacting** the data as needed, per BP 23).

Technical Specifications + Resources

Refer to BP 6, *Technical Specifications + Resources*, 'Risk management resources'.

Legal Framework

See section 4.2.2.C. on the factors relevant to finding that an interference with the right to privacy is proportionate, and can therefore be seen as justified. In particular, see the discussion on limiting who has access to data.
See section 4.2.4.B on 'Scope of the GDPR', 'Data Protection Measures', and 'Data Transfers' under the GDPR.
See also the European Court of Human Rights decision in *Big Brother Watch and Others v. United Kingdom*, Judgment, para. 362, wherein the Court stated that, in relation to data transfers, 'the transferring [entity] must ensure that the receiving [entity], in handling the data, has in place safeguards capable of preventing abuse and disproportionate interference. In particular, the receiving [entity] must guarantee the secure storage of the material and restrict its onward disclosure.'

APPLICABLE ETHICAL PRINCIPLES: Do No Harm; Legal Awareness; Accountability; Accuracy, Impartiality, and Objectivity.

[31] GDPR, Article 46.
[32] GDPR, Article 46; *Big Brother Watch and Others v. United Kingdom* (ECtHR), Judgment, para 362.

21. Obtain the consent of the data subject(s) prior to audio data transfer, if practicable.

Prior to the transfer of **audio data** to a **third-party recipient**, the **Collector** should take steps to obtain the audio **data subject**'s written or recorded **informed consent** with regard to the forthcoming transfer in accordance. This may only be practicable if the Collector was able to obtain consent from the data subject(s) at the point of collection of the data and maintained an avenue for later communication.

The data subject(s) in question may also be the Collector's **personnel**, for example if personnel are sending contemporaneous voice note updates from on the ground in a conflict context. The Collector should obtain the involved personnel's consent prior to any transfer of the data.

If the data subject denies consent to the transfer of the **audio data**, the denial should be documented and the Collector should not share the data, unless the data can be **anonymised** (for example, by **redacting** the data as needed, per BP 23). If it is not possible to obtain the data subject's consent prior to transfer of the data, the Collector should assess whether the data should still be transferred per the ethical principles outlined in BP 1.

Technical Specifications + Resources

Refer to BP 1 *Technical Specifications + Resources*, 'Resources on Consent'.

Legal Framework

See section 4.2.4.B on 'Legal Bases for Data Processing' under the GDPR.

APPLICABLE ETHICAL PRINCIPLES: Consent.

22. Enter into a data transfer agreement with relevant third parties prior to transferring audio data.

A data transfer agreement with a **third-party recipient** should clarify the nature of the data transfer arrangement between the **Collector** and the third party and the respective roles and responsibilities of the parties.

The data transfer agreement should be in writing (see BP 3), i.e. in the form of a contract or a memorandum of understanding. The agreement must stipulate the specifically designated purpose(s) for which the **audio data** may be used by the third party and the third party's commitment to adhere to said purpose(s).[33] The agreement may also indicate the parameters of the transfer relationship, for example, whether the information sharing will be proactive (whereby the Collector agrees to send all collected and relevant audio data to the third-party recipient) or reactive (requiring the third-party recipient to request the data from the

[33] Ljubljana-The Hague Convention, Article 16(1). Additionally, this practice is an extension of the purpose limitation requirement under the right to privacy and data protection, *see* ECtHR, Guide to the Case-Law of the European Court of Human Rights: Data Protection (2022), paras. 117-120; *Digital Rights Ireland v Minister for Communications, Marine and Natural Resources and others* (CJEU), Judgment, para. 61; GDPR, Article 5(1)(b) and Recital 50.

Collector, and potentially also specify the parameters requested, e.g., all data collected concerning a certain timeframe and/or location). Parameters of the agreement may additionally include transfer procedures, such as including an itemised list of the data files in the transfer to be approved and signed by the sender prior to transfer and by the recipient upon receipt.

The data transfer agreement should be drafted and signed with the advice of competent legal representation. It should be signed by the Collector and the relevant third party before any information sharing takes place, and securely stored (see BP 6).

Technical Specifications + Resources

Refer to BP 6, *Technical Specifications + Resources*, 'Data Transfer'.

For an indication of what information may be helpful to document regarding the transfer of data from one actor to another, see e.g., The Folke Bernadotte Academy and The Swedish National Defence College, A Handbook on Assisting International Criminal Investigations (2011), page 57.

On the desirability of crafting data-sharing policies, see e.g., OHCHR, Berkeley Protocol, page 49.

Legal Framework

See section 4.2.4.B on 'Data Transfers' under the GDPR.
See also Ljubljana-The Hague Convention, Article 16 on 'Use and protection of personal data'.

APPLICABLE ETHICAL PRINCIPLES: Accountability; Accuracy, Impartiality, and Objectivity.

23. If necessary, redact the duplicate audio data and its metadata before sharing.

Prior to transferring **audio data** to a **third-party recipient**, it may be necessary for portions of the audio data to be **redacted** for security or privacy purposes, e.g., to **anonymise** the **audio** with regard to the audio **data subject**(s) or any other **personal data** included (see BP 5).

Redaction should be handled in the same way as **enhancement** of audio data (per BP 15), meaning any redaction of audio data must be performed on a **duplicate copy**—not on the **original copy**—and should be duly documented in accordance with BP 3.

The **metadata** of any redacted duplicate copy should be similarly redacted. For example, the respective transcript and, where applicable, translation(s), should mirror the audio redaction with a visual redaction, for example, with a black rectangle, or by replacing the relevant word(s) with '[REDACTED]'.

The **Collector** should assess the necessity of redactions prior to the transfer of audio data in accordance with the data transfer risk assessment (see BP 20). The Collector should additionally assess the extent to which, in the event the audio is tendered as **evidence**, the redaction of names and other identifying information might affect the fair trial rights of an accused who typically has the right to such information.

Technical Specifications + Resources

Tools to anonymise data include, e.g., noise addition; speech transformation; voice conversion. Refer to BP 5 *Technical Specifications + Resources*.

An example of an approach to cryptography is zero-knowledge proofs, which are a class of cryptographic protocols that offer a way to verify secret information while keeping the information itself otherwise hidden. See e.g., K. Bamberger et al, Verification Dilemmas in Law and the Promise of Zero-Knowledge Proofs (2022), which provides both a technical primer and case studies about verification/authentication of the zero-knowledge mathematical demonstrations in legal spaces.

Legal Framework

See section 4.2.2. on the three-part test for determining whether an interference with the right to privacy is justified, and in particular section 4.2.2.B. on the necessity aspect of this test.
See section 4.2.4.B on 'Data Protection Measures' under the GDPR.
See section 4.3. detailing the rights of accused persons in connection with the right to a fair trial, in particular the right to be given exculpatory material.

APPLICABLE ETHICAL PRINCIPLES: Do No Harm

Best Practices for Specific Audio Data Collection Methods

Whereas the Best Practices until now apply equally to a variety of **audio data** collection methods, those included in this section are specific to certain audio data collection methods. This section includes BPs specific to audio download, audio **scraping**, and radio interception.

Audio Data Collected via Download

In this context, the term 'download' refers to the situation where a **Collector** finds **audio data** on, for example, a website (e.g., X/Twitter), an instant messaging app (e.g., Telegram), or in an email, and downloads the data to the Collector's computer system.

24. Collect information that contextualises the downloaded data

Where available, **Collectors** must collect information that speaks to the date, circumstances, and context of the downloaded **audio data**.[34] Doing so will improve the data's potential **evidentiary value** by helping to establish the data as **reliable** and **authentic**.

[34] *Prosecutor v. Bemba* (ICC), Public Redacted Version of "Decision on the Prosecution's Application for Admission of Materials into Evidence Pursuant to Article 64(9) of the Rome Statute" of 6 September 2012, para 84; *Prosecutor v Katanga and Chui* (ICC), Decision on Bar Table Motion, para 24; *Prosecutor v Ntaganda* (ICC), Decision on Prosecution's request for admission of documentary evidence, para 68.

In deciding what information to collect alongside the downloaded audio data, Collectors should consider guidance from the Berkeley Protocol, which identifies the following categories of information as being important:[35]

- Target web address: in the event that the audio data is downloaded from a web page, such as a social media site, the uniform resource locator (URL) should be recorded.
- Source code: in the event that the audio data is downloaded from a web page, Collectors should capture the HTML source code of the web page, if applicable.
- Full-page capture: Collectors should take a screen capture of the web page, messaging app, or email service to record what could be seen at the moment of collection. The date and time of the screen capture should be recorded.
- Additional media files: if the web page, messaging app, or email from which the **audio** is downloaded contain additional media files, such as videos or photos, Collectors should consider downloading these files. This decision should be made in light of the principle of data minimisation (see BP 2).
- Contextual data, including **embedded** and **custom metadata**: where available, data which provides context about the downloaded audio data should be collected, such as upload and/or author information, date and time, and comments and **tags**, and previously created **hash values**.
- Collection data: Collectors should, in accordance with BP 3, record the process of collecting audio data in careful detail, including the name of the person carrying out the collection, the IP address of the machine used in the collection, the virtual identity used, if any, and a time stamp.

Further to the above, where available, Collectors must collect information about the author of the audio data.[36] Where the author's identity is not immediately available, Collectors should consider the feasibility of investigating further to establish authorship, including any potential security risks such an investigation could pose.

Legal Framework

See section 5.3. on the role of date, circumstances, context, and authorship in establishing the probative value of data.

[35] OHCHR, Berkeley Protocol, page 59. Note that the categories and content have been adapted to reflect the Hala Protocol's focus on audio data originating from both open and closed sources.

[36] *Prosecutor v. Renzaho* (ICTR), Decision on Exclusion of Testimony and Admission of Exhibit, paras 1-2; *Prosecutor v Delalic* (ICTY), Decision on Admissibility of Evidence, paras 20-22, cited with approval in *Prosecutor v Brdjanin and Talić* (ICTY), Order on the Standards Governing the Admission of Evidence, para 18.

Audio Data Collected via Data Scraping

BPs 25-27 address **audio data** collected via data **scraping**, which is a process of automatically extracting data from a website by using a web scraper to access a web page and interpret and extract data.

25. Do no harm when data scraping

Data **scraping** has the potential to be very computationally demanding, which can place a burden on the server that is subject to the scraping activity. Data scraping can overwhelm servers if not done responsibly, rendering the server and the data it contains unavailable. **Collectors** should therefore carefully consider how they carry out their data scraping activities in order to do no harm to the servers or data involved.

Legal Framework

APPLICABLE ETHICAL PRINCIPLES*:* Do No Harm

26. Programme the data scraping tool with data minimisation in mind

Data **scraping** is a collection method that is capable of automatically collecting large volumes of data in a short space of time. As such, if the **collection effort** falls under the purview of the GDPR, the **Collector** must implement the principle of data minimisation in the programming of the data scraping tool (see BP 2).[37] This involves instructing the data scraper to collect only that data which is necessary to achieve the aims of the collection effort, for example data relating to a particular date or place.

If the GDPR does not apply to the collection effort, Collectors should still adhere to data minimisation as a best practice as this increases the chances that any infringement on privacy in the data collection effort will be considered proportionate and therefore justified. The automated characteristics of data scraping makes this particularly advisable, as privacy protections tend to be more stringent when applied to automated processes.

Legal Framework

See section 4.2.2.C. on the role that the principle of data minimisation plays in protecting the right to privacy. Observing data minimisation increases the chances that a data collection effort's infringement on privacy will be considered proportionate and therefore justified.
See also *Big Brother Watch and Others v. United Kingdom*, Judgment, para. 330, where the Court found that 'the need for safeguards will be all the greater where the protection of personal data undergoing automatic processing is concerned'.
See section 4.2.4.B on 'Scope of the GDPR' and 'Data Protection Measures'.

APPLICABLE ETHICAL PRINCIPLES: Legal Awareness

[37] GDPR, Article 5(1)(c).

27. Consider the feasibility of notifying the data subjects whose data was collected through scraping

Collectors falling under the purview of the GDPR or Convention 108 must notify **data subjects** of the collection of their personal data unless they fall under the exemption provided for by Article 14 of the GDPR and Article 8(3) of Convention 108, described below.[38] Collectors not covered by either instrument should still consider whether they fall under the terms of the exemption, and if not, consider notifying the data subjects. Notification can entail, for example, the Collector contacting the individual(s) whose personal data has been collected via email or post.

To qualify for the exemption, Collectors must demonstrate that:

- notifying the data subject about the collection of their personal data would be impossible; or
- notifying the data subject about the collection of their personal data would involve disproportionate effort; or
- notifying the data subject about the collection of their personal data would make achieving the **collection efforts** impossible or seriously impair them.[39] (This exemption is particular to the GDPR and is not included in Convention 108)

The exemption in Article 14 of the GDPR comes with a number of procedural safeguards that Collectors must comply with. As such, Collectors to whom the GDPR applies are advised to seek specialist advice on qualifying for the Article 14 exemption and the procedures to follow for compliance with the exemption conditions.

Legal Framework

See section 4.2.4.B. on 'Scope of the GDPR' and 'Notification Requirements' under the GDPR.
See section 4.2.4.A. on Convention 108.

APPLICABLE ETHICAL PRINCIPLES: Legal Awareness

Audio Data Collected via Radio Interception

BPs 28 and 29 concern data collected by intercepting radio signals. Radio interception involves the use of relevant equipment to make the content of a radio communication available to a person who is not the sender or intended recipient of that communication.

28. Place radio interception equipment in areas that pose minimal risk to the civilian population.

Radio interception equipment comprises the hardware and software necessary to intercept and collect radio signals, and can include an antenna, a receiver, and a computer. Radio interception requires that

[38] GDPR, Article 14; Convention 108, Article 8(3).
[39] GDPR, Article 14(5)(b); Convention 108, Article 8(3).

components of the equipment be placed in geographical proximity to the location from which the signal is transmitted.

The placing of radio interception equipment in civilian areas can create a security risk. Military and other actors seeking to prevent the collection of radio signals may target areas where interception equipment is known to be positioned, increasing the possibility of damage to civilian buildings and injury and loss of life. **Collectors** should therefore give due consideration to this risk in accordance with BP 6. Namely, prior to placing radio interception equipment on the ground, Collectors should carry out a risk assessment to identify any risks their collection activities could pose to the civilian population in the area. If risks are identified but, for operational reasons, interception equipment needs to be placed in that area, steps should be taken to minimise the risk to civilians, for example by placing units in abandoned buildings rather than inhabited ones. The process of determination of risk should be documented in accordance with BP 3.

Legal Framework

See section 4.1. on the duties applicable to private entities under the OHCHR, UN Guiding Principles on Business and Human Rights, and additionally for actors operating in conflict contexts, under the OHCHR, UN Guide on Heightened Human Rights Due Diligence for Business in Conflict-Affected Contexts.

APPLICABLE ETHICAL PRINCIPLES: Do No Harm

29. Demodulate radio signals that may contain relevant information.

Any **audio data** that is characterised as **signal** and which may include **relevant** information must be demodulated.[40]

Where the demodulation process leads to intelligible audio—for example, a voice or a possible gunshot—this audio can then be assessed for relevance per BP 13. If the audio remains unintelligible or inaudible following demodulation, or if the signal clearly only contains static or silence, it may be **deleted**. However, the Collector should consider still **preserving** the signal in case it could be rendered intelligible by future technology.

The Collector should also consider the feasibility of storing and preserving the non-demodulated signal (the 'raw' signal) within, or linked to, the **audio data file**.

Legal Framework

See section 5.2. on the importance of evidence being intelligible in order to be considered relevant.

[40] This is an analogical application of the case law of international criminal courts and tribunals concerning call data records—a form of metadata that provides information about the communication, including source, date, time and duration of the call. Call data records have been considered inadmissible on the basis of their unintelligibility. See *Prosecution v Ayyash et al.* (STL), Judgment, paras. 375–378, where the Trial Chamber rejected the admission of call data records due to it being voluminous and unreadable and containing a string of numbers and symbols.

PART 2: LEGAL FRAMEWORK

1. Introduction

The Best Practices for the Collection, Processing, and Transfer of Audio Data are informed and shaped by legal rules and ethical principles. The Legal Framework aims to identify the legal rules and, where relevant, indicate how they might apply to organisations working with audio data. The legal rules derive from international and regional legal frameworks.

The Legal Framework starts by discussing the relevance for the Protocol of the distinction between open- and closed-source information (Section 2). Section 3 provides a snapshot of the legal frameworks that organisations should consider when working with audio files, i.e. public international law, international human rights law, international humanitarian law, international criminal law, and domestic law. Sections 4 and 5 engage with the regional and international legal frameworks in greater detail. Section 4 discusses applicable international and regional European human rights law—particularly the rights to privacy and fair trial; Section 5 concerns international criminal law and discusses key evidentiary concepts.

2. A Preliminary Point: Open-Source vs. Closed-Source Audio Data

The discourse on digital information and evidence often centres around the distinction between open- and closed-source data. This section defines these concepts and explains their relevance for the Protocol.

The Berkeley Protocol on Digital Open Source Investigations (Berkeley Protocol) defines open-source information as 'publicly available information that any member of the public can observe, purchase, or request without special status or unauthorized access'.[41] In other words, open-source information is that which anyone can access without breaching legal, privacy, or security controls. While the threshold of what constitutes open source is not clearly defined, information that is public but challenging to access (such as that requiring technical skills or specialised software) may still be considered open-source so long as it is not obtained via unauthorised access. Examples of what would constitute unauthorised access include tapping a person's mobile phone to listen to their conversations or accessing their private email to read their messages without consent or a search warrant issued by a court.

Closed-source information is 'information with restricted access or access that is protected by law'.[42] This would include mobile phones and private emails, as well as information requiring official clearance to access. It would further include audio messages sent over instant messaging apps such as WhatsApp or posted to (semi)private channels on apps such as Telegram. It is possible to legally obtain closed-source information through non-public channels, such as judicial processes, or if the owner of the information voluntarily discloses it. However, what distinguishes it from open-source information is that it is only available to some members of the public.[43]

[41] OHCHR, Berkeley Protocol on Digital Open Source Investigations: A Practical Guide on the Effective Use of Digital Open Source Information in Investigating Violations of International Criminal, Human Rights and Humanitarian Law (2022) (OHCHR, Berkeley Protocol), page 6.
[42] OHCHR, Berkeley Protocol, page 6.
[43] OHCHR, Berkeley Protocol, page 6.

Despite the prominence of the distinction between open and closed sources in the discourse, the research carried out for the Hala Protocol indicates that, when it comes to establishing the evidentiary value of audio data, the open or closed-source nature of the data is not decisive. For example, as noted in section 5.1., the unlawful recording of closed-source phone calls did not affect the admissibility of the recordings as evidence before an international criminal tribunal. The research indicates that measures to safeguard the authenticity and verifiability of the data are more important than the type of source the data comes from.

The open or closed-source nature of data can have an effect on the right to privacy, which in turn, and in limited circumstances, can affect the admissibility of evidence. This possibility is covered in sections 4.2.1., 4.2.2. and 5.1. below. Furthermore, where a Collector is operating in a way that violates the right to privacy in a serious and systematic way, this can jeopardise their relationship with accountability mechanisms because the Collector's conduct will be considered unethical. This limits the potential for the collected audio data to be used for accountability.

The above findings are based on international case law and procedural rules before international criminal courts and tribunals. The relevance of the open or closed source-nature of the data may be greater at the domestic level depending on the legal rules applicable in individual jurisdictions. For this reason, Collectors should obtain expert legal advice from a lawyer qualified in the relevant jurisdiction.

3. An Overview of Applicable Legal Frameworks

This section provides an overview of the legal frameworks that organisations working with audio data should be aware of.

3.1. Public International Law

Public international law (PIL) is a vast body of law that governs relations between States and between States and private actors (e.g., individuals). PIL contains several branches of law that are relevant to work with audio data, such as international human rights law (IHRL), international humanitarian law (IHL), and international criminal law (ICL). Each is discussed below.

3.2. International Human Rights Law

IHRL is a branch of PIL that sets obligations and duties on States to respect, protect, and fulfil the human rights of individuals. The commitment to respect requires States to refrain from interfering with or curtailing the enjoyment of human rights. The obligation to protect demands that States protect individuals and groups against human rights violations perpetrated by others (including State authorities, individuals, or other private actors such as corporations). Finally, the obligation to fulfil requires States to take positive action to facilitate and ensure that the individuals within their jurisdiction fully enjoy their human rights. These obligations are generally set out in IHRL treaties ratified by States.[44] The rights outlined in IHRL

[44] The Universal Declaration of Human Rights 1948 (UDHR), combined with the International Covenant on Civil and Political Rights 1966 (ICCPR), and the International Covenant on Economic, Social and Cultural Rights 1966, establish a foundation of defined human rights known as the International Bill of Rights. Other IHRL treaties include, e.g., the International Convention on the Elimination of All Forms of Racial Discrimination 1965, the Convention on the Elimination of All Forms of Discrimination against Women 1979, the Convention against Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment 1984, and the Convention on the Rights of the Child 1989.

treaties include the right to life, the right to liberty, the right to a fair trial, and the right to privacy. PIL also includes regional human rights law instruments,[45] some of which are discussed below in section 4.

3.3. International Humanitarian Law

IHL is a branch of PIL that regulates how States and non-State actors should behave during an armed conflict. IHL does not relate to whether an armed conflict is lawful or not. Rather, it dictates what conduct is permitted or prohibited in the context of an armed conflict once it has broken out. For example, IHL prohibits the targeting of civilians, the use of certain weapons, and contains standards on how the warring parties should treat prisoners of war. IHL applies to armed conflicts between one or more States as well as conflicts involving States and non-State actors (such as paramilitary groups). Non-State actors are also obliged to respect the rules of IHL.

In addition to the above, IHL contains rules that identify when a civilian can be lawfully targeted. In some cases, when a civilian is taking part in hostilities, for example by firing weapons at the opposing side or by delivering military intelligence, they may be targeted as if they were a member of the armed forces. Protection is lost for the time that the participation in hostilities goes on, but resumes when participation ceases.[46]

3.4. International Criminal Law

ICL is a branch of PIL that applies to individuals, rather than States, and identifies when an individual's conduct will amount to an international crime. When an international crime has been committed, ICL sets out a framework for holding perpetrators to account. There are four core crimes under ICL: genocide, crimes against humanity, war crimes, and the crime of aggression.[47] These crimes are distinct from ordinary crimes in the sense that they are characterised by specific contextual elements, which are a set of factors that must be additionally satisfied before such a crime can be said to have been committed. The war crime of wilful killing, for example, is characterised by the fact that not only must there be an intentional killing, but it must be committed in the context of and in association with an armed conflict.

3.5. Interaction between IHL, IHRL, and ICL

In situations of armed conflict, IHL, IHRL, and ICL apply concurrently. The connection between IHL and ICL stems from the fact that serious violations of IHL can also constitute war crimes under ICL. For example, the killing of wounded soldiers who have laid down their arms is a violation of IHL. If the violation is committed by an individual with the necessary knowledge and intent, this can constitute a war crime.

The relationship between IHL and IHRL is more complex because contradictions can arise between the two legal frameworks. For example, IHL permits the killing of soldiers and, under some limited circumstances, of civilians. In contrast, a cornerstone of IHRL is the right to life. It is not the case that IHRL ceases to apply during armed conflict. Instead, the two branches of law apply at the same time. Determining which rules

[45] For example, key European human rights instruments are the Council of Europe's Convention for the Protection of Human Rights and Fundamental Freedoms 1950 (ECHR) and the European Union's Charter of Fundamental Rights 2012; the key human rights instrument of the Organization of American States is the American Convention on Human Rights 1969 (ACHR); the key human rights instrument of the African Union is the African Charter on Human and People's Rights 1981 (ACHPR).
[46] *See* Nils Melzer, 'Interpretive Guidance on the Notion of Direct Participation in Hostilities under International Humanitarian Law' (2008) 90 Intl Rev of the Red Cross 991; *See also* ICRC, Direct Participation in Hostilities: Questions and Answers (2009).
[47] ICC, Rome Statute of the International Criminal Court 1998 (Rome Statute), Articles 6, 7, 8 and 8*bis*.

take precedence in a given situation depends on the particular circumstances: for example, IHL would take precedence and apply to conduct during a battle/military operation, whereas IHRL would apply to a civilian protest (even if this protest takes place in a State experiencing armed conflict).

3.6. Domestic Law

A State's domestic legal system regulates the conduct of State actors, natural persons, and legal persons within its territorial borders. On the one hand, domestic law might incorporate provisions of international law, especially of IHRL and ICL. This could be done, for example, by the criminalisation of certain conduct under domestic law. On the other hand, a State's domestic law might differ somewhat from international law. For example, if an international treaty is not universally ratified or well-enforced, there may be States with domestic laws that conflict with the principles outlined in the treaty.

4. International Human Rights Law (IHRL)

The Best Practices identified in this Protocol reflect standards found in IHRL. Collectors may question why they should work to IHRL standards when, as explained in Section 3.2. above, IHRL is aimed at States. Collectors, as private entities, should operate in a manner that respects IHRL standards for several reasons:

- The domestic law of the State where the audio data collection takes place (and where the processing takes place, if these are different) will often require private entities to conform with international and regional human rights law;
- Private entities should act in accordance with the UN Guiding Principles on Business and Human Rights (UNGPs), which require business enterprises to respect human rights.[48] While the UNGPs are not legally binding, and while they are not targeted at non-commercial entities *per se*, they are the authoritative global standard and frequently used as a basis for domestic and regional legal measures that are binding on private entities;
- Audio data will be considered more robust before a court if it has been collected in compliance with IHRL standards;[49] and
- Operating within a human rights protective framework helps to uphold the integrity of a Collector's work and its reputation among third parties.

[48] OHCHR, Guiding Principles on Business and Human Rights, 2011 (OHCHR, Guiding Principles on Business and Human Rights), page 13.
[49] *See infra*, Section 5 below.

A Note on Derogations

It should be noted that some IHRL treaties contain derogation provisions that, subject to strict requirements,[50] allow States to temporarily suspend some of their human rights obligations[51] in order to take measures to protect the security of the State in times of war or when a public emergency threatens the life of the nation. Collectors tend to operate in States that experience frequent or ongoing armed conflict, and, as a result, they may have officially derogated from some human rights obligations. A derogation by a State should not alter a Collector's practices and the human rights law standards they aim to uphold; they should continue to operate as if there were no derogation. This is important to ensure that the Collector's operations are aligned with international human rights standards regardless of the domestic context in which they operate.

4.1. UN Guiding Principles on Business and Human Rights

The UNGPs[52] are guidelines requiring companies to respect human rights and provide a remedy for business-related human rights violations that companies may have played a part in. The UN Guide on Heightened Human Rights Due Diligence for Business in Conflict-Affected Contexts (hHRDD Guide) offers more specific guidance on how the UNGPs apply during conflict.[53] The UNGPs require that companies identify, prevent, mitigate, and account for the adverse social and human rights impacts of their business activities. CSOs are not companies and are, therefore, not directly addressed by the UNGPs. However, the elements of the UNGPs identified in this section are highly relevant for CSOs and are easily transferable to their activities. The UNGPs should, therefore, be considered as part of the legal landscape that CSOs operate in.

Collectors operating in conflict-affected areas, whether they be businesses or CSOs, have heightened human rights duties because their activities can impact the dynamics of a conflict. Firstly, Collectors must ensure that their personnel do not perpetrate human rights abuses; secondly, Collectors must take steps to avoid enabling, exacerbating, or facilitating a serious human rights abuse by virtue of its activities. The duties contained in the UNGPs and the hHRDD Guide can be operationalised as follows:

- *Do not make a crime possible*: Collectors should not provide a person or group perpetrating (or at risk of perpetrating) a crime with materials that could assist in a violation of IHRL or IHL. A Collector could be enabling a crime if it provides materials such as weapons, vehicles, fuel, or information.
- *Do not make a crime easier to carry out*: Collectors should ensure that they do not make the commission of a crime easier. For example, if a company provides sophisticated tracking software to an armed group, it increases the group's efficiency in targeting specific individuals. If targeting is then done for criminal purposes, the company can be held criminally responsible.

[50] The ICCPR, Article 4 and ECHR, Article 15 both permit member states to derogate from their obligations 'to the extent strictly required by the exigencies of the situation, provided that such measures are not inconsistent with their other obligations under international law.'
[51] Not every right can be derogated from. For example, the prohibitions on torture and slavery cannot be limited in any way.
[52] OHCHR, Guiding Principles on Business and Human Rights.
[53] UNDP, Heightened Human Rights Due Diligence for Business in Conflict-Affected Contexts: A Guide (2022) (UNDP, hHRDD Guide); Jonathan Kolieb, 'Don't forget the Geneva Conventions: achieving responsible business conduct in conflict-affected areas through adherence to international humanitarian law' (2020) 26 Australian Journal of Human Rights 142.

- *Do not make a crime worse*: it is important that a Collector not increase the gravity of a crime by contributing to it. For example, a business may aid and abet a crime by selling a product that will increase the gravity of an attack against a group of civilians.

Criminal liability may follow if it can be proven that a company is responsible for such conduct and knew or should have known that its actions would contribute to human rights abuses or violations of IHL. Accordingly, Collectors should maintain and regularly update policies and practices that align with all relevant principles while operating in a conflict-affected area.

To identify any potential negative impacts that a Collector may have on persons and groups in the course of its operations in the context of a conflict, Collectors should conduct a human rights due diligence impact assessment prior to their operations in the area. Such a due diligence assessment should clearly identify the prevailing human rights situation in the area where the Collector plans to operate. In the simplest terms, this assessment may be done by asking the following questions:

- Is there an actual or potential adverse impact on human rights or the conflict connected to the company's activities (actions or omissions), products, or services in any of the State parties to the conflict?
- If so, do the company's activities in any of the State parties to the conflict increase the risk of that impact?
- If so, would the company's activities in any of the State parties to the conflict in and of themselves be sufficient to result in that impact?[54]

After the assessment, appropriate steps should be taken if any potential human rights risks are identified.

Collectors should periodically conduct such assessments in light of certain developments in their activities, including:

- Before a new activity or relationship;
- Before significant decisions or changes in the operation (e.g., market entry, product launch, policy change, or broader changes to the activities);
- In response to or in anticipation of changes in the operating environment (e.g., rising social tensions); and
- Periodically throughout the life of an activity or relationship.[55]

4.2. The Right to Privacy and Data Protection

In collecting audio data, collectors must consider the right to privacy. This right protects an individual's private sphere (private life, home, and correspondence) from interference by others. While privacy is a core human right, it is not an absolute right and can be derogated from and limited under certain circumstances. Derogation was addressed above in the introduction to section 4; limitations will be addressed in this section. **A right is limited when there is a justified interference with that right**.

[54] UNDP, hHRDD Guide, page 28.
[55] UNDP, hHRDD Guide, page 20.

An important element of the right to privacy is data protection.[56] The processing of personal data can constitute an interference with privacy. Data protection is relevant when working with audio because recordings of a person's voice can constitute personal data,[57] and actions taken concerning that data (collection, storage, alteration, disclosure, and erasure)[58] constitute processing. As such, whenever audio data collected and otherwise processed contains human voices, data protection issues arise.

What is personal data?

Personal data is understood as information relating to an identified or identifiable individual.[59] Data can identify a person in any way,[60] whether *directly* (for example, through a name or identification number), or *indirectly* (for example, through an IP address).[61] As such, personal data includes information that, when linked with other information, could lead to the identification of a particular person, even though the information on its own would not be enough for an identification.[62] For example, cell phone location data may not in and of itself be enough to identify an individual, but when it is combined with other information, such as property tax records, it may point to a specific person. As a result, cell phone location data is personal data. If data is anonymised, meaning that any identification is irreversibly prevented, then it will no longer constitute personal data.[63]

Data does not need to be reviewed or deciphered, or a person be identified, for it to constitute personal data.[64] As long as it is possible to identify a person from the data, it is personal data. As such, whether audio data constitutes personal data depends on whether the speaker *could be* identified using any aspect of the data, rather than whether the speaker *has been* identified.

While the right to privacy is included in most core international human rights treaties,[65] the analysis in this section of the Legal Framework focuses on relevant provisions of European human rights instruments, including Article 8 of the European Convention on Human Rights (ECHR) and multiple provisions of the Council of Europe Convention No. 108 for the Protection of Individuals with Regard to Automatic Processing of Personal Data 1981 (Convention 108). In addition, this section covers relevant obligations under European Union (EU) law, particularly the EU's General Data Protection Regulation (GDPR).

[56] '[C]ontemporary data protection frameworks are, conceptually speaking, legislative substantiations of the right to privacy' (Robin Geiß and Henning Lahmann, 'Protection of Data in Armed Conflict' (2021) 97 International Law Studies 556, page 568).
[57] ECtHR, Guide to the Case-Law of the European Court of Human Rights: Data Protection (2022), paras. 8 and 50-53. Under some circumstances, voice recordings can even be considered 'sensitive data' and therefore subject to heightened protection (para. 24) or otherwise as a category of data of special concern (paras. 50-53).
[58] Convention for the Protection of Individuals with Regard to the Automatic Processing of Individual Data 1981 (Convention 108) Article 2(b). Additionally, the term 'processing' should be interpreted as including any use of the collected data to train AI/develop algorithms. Such processing should therefore adhere to the same data protection requirements as any other form of processing.
[59] Convention 108, Article 2(a); GDPR, Article 4(1).
[60] GDPR: Personal Data (Intersoft Consulting).
[61] ECtHR, Guide to the Case-Law of the European Court of Human Rights: Data Protection (2022), para. 8.
[62] What is Personal Data? (European Commission)
[63] European Commission Data Protection Working Party, Opinion 05/2014 on Anonymisation Techniques, adopted 10 April 2014, page 9, which clarifies: 'An effective anonymisation solution prevents all parties from singling out an individual in a dataset, from linking two records within a dataset (or between two separate datasets) and from inferring any information in such dataset.'
[64] Among the examples of 'personal data' listed in ECtHR, Guide to the Case-Law of the European Court of Human Rights: Data Protection (2022), para. 8, is that of 'electronic data seized in a law firm, even though it had not been deciphered, transcribed or officially attributed to their owners' (citing *Kırdök and Others v. Turkey,* Judgment, ECtHR, 14704/12, 3 December 2019 (*Kırdök and Others v. Turkey,* Judgment), para. 36). Extrapolating from this example, data that has the potential to identify a person is still personal data even if it has not been deciphered, translated, or attributed.
[65] The right to privacy is protected by major regional and international human rights treaties: ECHR, Article 8; ICCPR, Article 17; ACHR, Article 11; EU Charter of Fundamental Rights, Article 7 (privacy) and 8 (data protection).

Not all Collectors operate in jurisdictions within the Council of Europe or EU, and therefore, not all Collectors are bound to uphold the standards below. That being said, European privacy and data protection standards represent a high level of human rights protection that Collectors should nevertheless strive to maintain in their work. As stated in the UNESCO Guidelines for Judicial Actors on Privacy and Data Protection, the 'implementation of data protection in its most widespread aspect is nowadays represented in the European context by the General Data Protection Regulation (GDPR), which has served as a basis and inspiration for much subsequent legislation worldwide'.[66]

There are two key steps in assessing the right to privacy implications of collecting audio data: 1) determining whether the right to privacy is engaged, and 2) if it is, determining whether an interference with the right is justified.

4.2.1. *Whether the Right to Privacy is Engaged*

Audio data collected in connection with accountability work can differ in nature. For example, in a conflict context it may capture communications between military personnel, communications between civilians, and/or communications between military and civilians. It may capture open-source communications—such as those communicated over an open radio frequency—or closed-source communications—such as a telephone conversation or a voice message sent with an instant messaging app. The privacy concerns will differ depending on the kind of communication involved.

Collectors should note that the right to privacy under international human rights law continues to apply in an armed conflict.[67] Even if the State on whose territory the conflict is taking place has derogated from the right to privacy, Collectors should continue to uphold and respect the right to privacy in their work (see discussion on derogations above at the beginning of section 4).

A. Military Communications and the Right to Privacy

Military communications, understood as communications between military personnel, are made in the context of a state function, which is, by nature, a public function and not for the purpose of personal fulfilment or development. The ECtHR has found that activities of 'an essentially public nature' are outside of the scope of private life and fail to offer the respective actor a reasonable expectation of privacy.[68] The jurisprudence, therefore, supports the position that communications carried out for public purposes, such as in relation to military activity, do not involve a reasonable expectation of privacy.[69]

Military communications will be excluded from privacy protection regardless of whether the communications take place using open-source or closed-source channels. When the nature of the communication is public, the fact that it is transmitted through a closed-source channel cannot convert that public nature into a private one.

---

[66] UNESCO, Guidelines for Judicial Actors on Privacy and Data Protection (2022), page 17.
[67] Mary Ellen O'Connell, 'Data Privacy Rights: The Same in War and Peace' in Russel Buchan and Asaf Lubin (eds) *The Rights to Privacy and Data Protection in Times of Armed Conflict* (NATO CCDCOE, 2022), page 13.
[68] *Friend and Others v UK*, Decision, ECtHR, 16072/06 27809/08, 24 November 2009, para. 42.
[69] *Friend and Others v UK*, Decision, ECtHR, 16072/06 27809/08, 24 November 2009, para. 42. *See also*, *Uzun v Germany*, Judgment, ECtHR, 35623/05, 2 September 2010 (*Uzun v Germany*, Judgment), para. 44: a reasonable expectation of privacy is a significant, but not conclusive factor, when assessing whether a person's private life is concerned.

The application of privacy protections is more complex when communications by military personnel are not (purely) military in nature. For example, a member of the military may use a radio or mobile phone to call family members for personal reasons. While the ECtHR has not adjudicated on this specific scenario, the matter would most likely depend on the type of communication channel used. If members of the military use open-source communication channels for private communications, this would likely not be protected. ECtHR case law provides that a violation of privacy requires personal data to have been compiled, processed, or published in a way beyond that which is reasonably foreseeable.[70] Where military personnel communicate over open-source channels—for example open radio frequencies that are unencrypted and publicly available–there is a reasonable foreseeability that the communications will be accessed, heard, and widely collected and shared by interested parties and stakeholders, including the opposing side of a conflict. By contrast, military personnel could expect communications over closed-source channels–for example, mobile phones—to be less easily accessed. The expectation of privacy, therefore, would be higher (although not absolute, as in conflict contexts the opposing side can always be expected to have an interest in eavesdropping on military communications of all kinds).

Given the nature of modern day conflict, Collectors will often face situations where they collect the audio communications of non-State armed groups. While fighters within these groups are not members of a State military, the above analysis should apply in the same way. Non-State armed groups often exercise effective control over segments of territory and population, such that for the individuals under their control, they are the *de facto* public authority. Thus, the same considerations for reasonable expectation of privacy held by State Military can be applied to Members of non-State armed groups (where they carry out functions that would be classified as public if a member of a State military carried them out).

Where there is doubt as to whether the right to privacy protects a particular audio communication, Collectors should err on the side of caution and assume that it is protected. In that case, Collectors should assess whether the interference is justified (discussed in section 4.2.2).

B. Civilian Communications and the Right to Privacy

Audio data collection efforts may pick up civilian communications and military communications, either incidentally or by design. Given the potentially sensitive nature of voice recordings, it is best to assume that the right to privacy of a civilian whose voice has been recorded is engaged from the moment of collection. Civilian communications may be entirely unrelated to military activities, or they may touch on military activities in some way, such as when two civilians discuss the arrival of soldiers in their village. In both cases, privacy is engaged because the communications relate to personal development, including establishing relationships with others and the outside world.[71]

Even when they are not members of the military, it is possible for civilians to directly participate in hostilities 'when they carry out acts, which aim to support one party to the conflict by directly causing harm to another party, either directly inflicting death, injury or destruction, or by directly harming the enemy's military

[70] *Uzun v Germany*, Judgment, para. 45; *Oy and Oy v Finland*, Judgment, ECtHR, 931/13, 27 June 2017, para. 136. A significant element in determining whether operations concerning personal data fall within the scope of Article 8 is whether an individual is entitled to expect protection of his/her private life (ECtHR, Guide to the Case-Law of the European Court of Human Rights: Data Protection (2022), para. 13).
[71] ECtHR, Guide on Article 8 of the European Convention on Human Rights (2022), para. 79.

operations or capacity'.[72] The effect of direct participation is that civilians temporarily lose protection from being targeted under the rules of IHL as long as they carry out such acts. In such a scenario, it is arguable that the right to privacy would not cover communications because they are related to a public function in the same way as military communications.

4.2.2. *Whether Interference with an Individual's Right to Privacy is Justified*

If a person's personal data has been collected in circumstances where the right to privacy is engaged, there will always be an *interference* with the right to privacy, but there will not always be a *violation* of the right to privacy. It is possible for the interference with privacy to be justified, and if justified, the interference would be a permissible limitation of the right and would not constitute a violation. Article 8 of the ECHR and Article 11 of Convention 108 list the following criteria for justifying an interference with the right to privacy;

- It must be in accordance with the law;
- It must be necessary to achieve a legitimate aim, namely national security, public safety, the economic well-being of the country, the prevention of disorder, the investigation and prosecution of criminal offences, the protection of health or morals, or the protection of the rights and freedoms of others; and
- It must be proportionate to the aim sought.

For an interference with privacy to be justified, the above criteria should be met at the outset of the interference and throughout. As time passes and circumstances change, regular evaluation will be required to be sure the criteria continue to be met.[73]

As already mentioned, IHRL is only legally binding upon States. As such, the justification criteria are intended to be applied to a State's interference with the right to privacy. The assessment below nevertheless applies the justification requirements to the operations of CSOs to demonstrate how these Collectors can adhere to IHRL standards.

A. In Accordance with the Law

The requirement that measures interfering with the right to privacy be 'in accordance with the law' is designed to prevent State authorities from acting arbitrarily. For this requirement to be met, State authorities must have a basis in domestic law for their actions, and this domestic law must provide appropriate safeguards. A domestic court must often give state authorities specific authorisation before taking measures interfering with an individual's privacy.

The importance of this requirement for CSOs is not settled. Given that IHRL is addressed to States, and given that private entities do not generally have standing to apply for permission to collect closed-source data relating to other private entities, it is possible that this requirement does not apply. That being said,

---

[72] *See* ICRC, Direct Participation in Hostilities: Questions and Answers (2009); *also see* Russel Buchan and Nicholas Tsagourias, Ukranian 'IT Army': A Cyber Levée en Masse or Civilians Directly Participating in Hostilities? (EJIL:Talk, 9 March 2022).

[73] *P.N. v Germany*, Judgment, ECtHR, 74440/17, 11 June 2020, para 85; *Catt v United Kingdom,* Judgment, ECtHR, 43514/15, 24 January 2019, paras 119-120; *Big Brother Watch and Others v. UK*, Judgment, ECtHR, 58170/13, 62322/14 and 24960/15, 25 May 2021 (*Big Brother Watch and Others v. United Kingdom*, Judgment), paras 350, 356; *Case of S. and Marper v United Kingdom*, Judgment, ECtHR, 30562/04 and 30566/04, 4 December 2008, para. 119.

private entities are subjects of domestic law; their collection efforts should therefore respect provisions of domestic law that apply to them.

Despite the uncertainty surrounding this requirement, Collectors can consider two factors to inform their approach: the nature of the data and the context of the collection.

Nature of the Audio Data

Data collection from open sources, including personal data such as photographs, is not generally regarded as requiring a basis in law or specific authorisation. In the November 2022 decision in *Ukraine and The Netherlands v Russia*, the ECtHR relied on open-source data collected by both private and public entities–including photographs–to establish the admissibility of a claim. No mention was made of a legal basis or authorisation being required to collect the data, nor was the issue raised by any of the parties to the proceedings.[74]

By comparison, access to closed-source data will require some form of legal basis and authorisation. In the same November 2022 decision, the ECtHR also relied on intercepted telephone communications. In relation to this closed-source data, the court noted that the collection and sharing of intercepts was authorised by a domestic court and the domestic prosecutor's office.[75] This accords with the long line of ECtHR case law on the right to privacy and closed-source data.[76] Many legal systems prohibit–through criminal law, civil law, or both–the unauthorised collection of closed-source data by non-state parties.

Context

In situations of armed conflict, it may not be possible or desirable to request authorisation to collect audio data, as doing so may compromise the objective of the collection effort. In such cases, international criminal courts and tribunals have been flexible when approaching closed-source audio data evidence collected otherwise than in accordance with the law.[77] Therefore, lack of authorisation does not affect the evidentiary value of audio data from closed sources *per se*.

B. Necessary to Achieve a Legitimate Aim

Collectors working in the accountability space have a strong claim that their audio data collection is necessary to achieve a legitimate aim. Audio data collected in the context of an armed conflict or serious violence can play a role in establishing the truth of events: it captures contemporaneous information that could be used in future criminal trials or other accountability efforts. Furthermore, the ephemeral nature of digital information means it can be taken offline or deleted at any moment, making its fast collection and preservation crucial to accountability efforts. This is particularly true of intercepted audio: if this audio is not captured in the precise moment that it is communicated, there will be no record of the communication and valuable potential evidence of criminal acts will be lost.

---

[74] *Ukraine and The Netherlands v Russia*, Decision, ECtHR, 8019/16, 43800/14 and 28525/20, 30 November 2022 (*Ukraine and NL v Russia*, Decision), see for example paras 464, 472, 524, 620, 650.
[75] *Ukraine and NL v Russia*, Decision, paras 1501-2.
[76] ECtHR, Guide to the Case-Law of the European Court of Human Rights: Data Protection (2022), paras 86-87
[77] *Prosecutor v Brdjanin*, Decision on the Defence "Objection to Intercept Evidence", Case No. IT-99-36-T, ICTY, 3 October 2003, (*Prosecutor v Brdjanin,* Decision on the Defence "Objection to Intercept Evidence"), para. 56; *Prosecutor v. Renzaho*, Decision on Exclusion of Testimony and Admission of Exhibit, ICTR-97-31-T, 20 March 2007 (*Prosecutor v Renzaho,* Decision on Exclusion of Testimony and Admission of Exhibit), para. 15.

At the necessity stage of the privacy assessment, an important question to ask is whether the legitimate aim could be achieved by less intrusive means. If the answer is 'yes', then the audio collection is not *necessary* to achieve the aim and less intrusive means should be sought.

C. Proportionate to the Aim Sought

The proportionality requirement for justifying an interference with privacy calls for the greatest nuance. An assessment must be made to measure whether the infringement of privacy is proportionate to the legitimate aim—in this case, of contributing to accountability for international crimes. This section identifies the considerations relevant to assessing proportionality in the context of work with audio data by formulating key questions that Collectors can ask themselves about their data practices. This is followed by a more detailed discussion of these issues.

Key Questions

Was the audio data open source or closed source?	There is generally a greater expectation of privacy in closed-source data, making proportionality harder to establish.
Did the person/s know they were being recorded?	If the subject of the audio recording was aware that the recording was being made, but nevertheless continued to speak, proportionality may be easier to establish.
Does the audio recording contain sensitive or privileged information?	Sensitive or privileged information is subject to higher protection, so its collection and processing will generally be harder to establish as proportionate.
How serious are the crimes to which the audio recording relates?	The more serious the crime, the lower the threshold for proportionality.
How long will the audio data be retained for?	Generally speaking, the longer the Collector plans to retain the audio data, the harder it will be to argue that the retention is proportionate.
Was the audio data collected for a clearly articulated and limited purpose?	The collection and processing of audio data will be easier to establish as proportionate if done for a concrete and clearly articulated purpose.
Is the audio data only being used for the purpose for which it was collected?	Where data collected for one purpose is later used for a different purpose, the proportionality of this secondary use of the data will be more challenging to establish.
Is data minimisation being observed?	The amount of data collected should be proportionate to the purpose for which it is collected and excessive collection should be avoided.
Are limits placed on who can access the data?	Access to the data should be limited to those who require access in order to achieve the purposes behind the collection. Unjustifiably broad access will be disproportionate.

Is the audio data linked with other personal data?	When different data points are linked together, this can provide a more complete picture of an individual's private life, making proportionality harder to establish.

Detailed Discussion

As a starting point, it is worth noting that there are no absolutes when it comes to assessing the proportionality of a privacy interference. It is entirely dependent on the constellation of factors in a given case. Collectors should therefore reassess the proportionality of their practices whenever something in their collection and/or processing changes.

In cases where the threshold for establishing proportionality is high, it does not follow that data collection and/or processing can never be proportionate. Proportionality can still be achieved by having strong protections in place. For example, if data needs to be retained for a long time, the retention can be proportionate if the crime to which the data relates is very serious and access to the data is strictly controlled.

Open-Source vs Closed-Source Data
The data's open source or closed source nature is relevant for proportionality because it affects a person's expectation of privacy regarding their data. While there is no direct ICL or IHRL case law on this point, one can analogise from the case law of the ECtHR concerning surveillance in private and public places.[78]

Collectors should work from the starting point that open-source data will have a low expectation of privacy attached to it, and closed-source data will have a high expectation of privacy. From this starting point, Collectors should also consider the nature of the data (whether it is particularly sensitive) and who made the data public (whether it was the data subject themselves or a third party[79]).

Knowledge of the recording
When collecting data relevant to the commission of international crimes, it cannot be a requirement that the data subject knew their voice was being recorded. Indeed, this could undermine evidence collection efforts. That being said, case law from the ECtHR indicates that it will work in favour of a finding of proportionality if the data subject was aware that they were being monitored.[80] In situations where individuals are speaking on open lines–such as unencrypted radio channels–or where they know their phone has been tapped,[81] the threshold for proportionality will be lower.

Sensitive or privileged information
Sensitive information, such as that pertaining to a person's religion, race, ethnicity, or sexual orientation, is subject to more stringent protection than other kinds of information.[82] In the view of the ECtHR, it is not acceptable to process this kind of data in line with 'ordinary domestic rules'.[83]

[78] ECtHR, Guide to the Case-Law of the European Court of Human Rights: Data Protection (2022), para. 158-161.
[79] *See* CoE, Guidelines on Safeguarding Privacy in the Media (2018), page 16.
[80] ECtHR, Guide to the Case-Law of the European Court of Human Rights: Data Protection (2022), paras. 153-157.
[81] In *Prosecutor v Brdjanin,* Decision on the Defence "Objection to Intercept Evidence", para. 63, the fact that the defendant knew their phone was tapped was important to finding that illegally obtained evidence could be admissible in a trial at the ICTY.
[82] GDPR, Article 9; Convention 108, Article 6; ECtHR, Guide to the Case-Law of the European Court of Human Rights: Data Protection (2022), paras. 18-37.
[83] ECtHR, Guide to the Case-Law of the European Court of Human Rights: Data Protection (2022), para. 21.

Privileged information is likewise more protected. It is in the public interest for communications between, for example, lawyers and clients, to be heavily protected from interference. The case law of the ECtHR leaves very little margin of appreciation to states to restrict the right to privacy in relation to privileged communications,[84] and the ICC Appeals Chamber in *Bemba et al* stressed that the recordings made of the defendants to which the Prosecution had access related only to non-privileged phone calls.[85]

The seriousness of the potential crimes

The more serious the crime, the easier it is to justify processing data that could aid in its investigation and prosecution.[86] As part of its assessment of proportionality in the context of the right to privacy, the ICC Appeals Chamber in *Bemba et al* indicated that the infringement should be proportionate to the investigative need.[87]

The period of data retention

Ideally, there should be a defined amount of time that data will be retained, after which it is deleted.[88] The longer data is retained, the harder it will be to argue that the interference with privacy is proportionate. Although it is not automatically a problem if there is no set end date for the retention of the data,[89] it can be an important consideration in a proportionality assessment. Regardless, a decision about the length of the data retention should be made on the basis of objective criteria and determined by necessity.[90]

In conflict situations, it can be many years before accountability processes materialise, making it impossible to know in advance how long data needs to be retained for.[91] Furthermore, it can be difficult to know beforehand what data may or may not be relevant to future accountability proceedings. If data is deleted that later turns out to be pertinent to the defence case, this can compromise the rights of the accused. With that said, deleting data once it is clear that the data is not relevant can support the proportionality of the retention period.[92]

Clearly articulated and limited purpose

The collection and processing of data for a specific purpose, which is clearly set out beforehand, is more likely to be proportionate than data collected for a vague and broad reason. Judges at the STL considered it relevant for proportionality that call data records were collected and stored to investigate 'concrete and specific crimes whose execution has already taken place' and not 'future indeterminate and unspecified criminal conduct'.[93]

---

[84] ECtHR, Guide on Article 8 of the European Convention on Human Rights (2022), para. 242.
[85] *Prosecutor v Bemba et al*, Decision on Requests to Declare Certain Materials Inadmissible, paras. 17-18.
[86] ECtHR, Guide to the Case-Law of the European Court of Human Rights: Data Protection (2022), paras. 199-202.
[87] *Prosecutor v Bemba et al.,* Judgment on the appeals of Mr Jean-Pierre Bemba Gombo, Mr Aimé Kilolo Musamba, Mr Jean-Jacques Mangenda Kabongo, Mr Fidèle Babala and Mr Narcisse Arido against the Decision of Trial Chamber VII entitled "Judgment pursuant to Article 74 of the Statute, ICC-01/05-01/13-2275-Red, 8 March 2018 (*Prosecutor v Bemba et al,* Appeal Judgment), para. 336.
[88] See ECtHR, Guide to the Case-Law of the European Court of Human Rights: Data Protection (2022), paras. 204-213. See also UN Security Council Counter-Terrorism Committee Executive Directorate, Guidelines to facilitate the use and admissibility as evidence in national criminal courts of information collected, handled, preserved and shared by the military to prosecute terrorist offences (2020) (UNSC CTCED, Military Evidence Guidelines), page 18, guideline 12.
[89] ECtHR, Guide to the Case-Law of the European Court of Human Rights: Data Protection (2022), para. 207.
[90] *Digital Rights Ireland v Minister for Communications, Marine and Natural Resources and others*, Judgment, CJEU, Joined Cases C-293/12 and C-594/12, 8 April 2014 (*Digital Rights Ireland*, Judgment), paras. 64 and 65.
[91] For example, crimes relating to the conflict in the former Yugoslavia are still being prosecuted, despite the war being over for many years and the first indictment of the ICTY being issued in 1994.
[92] ECtHR, Guide to the Case-Law of the European Court of Human Rights: Data Protection (2022), para. 205.
[93] *Prosecutor v Ayyash et al*, Decision on appeal by counsel for Mr Oneissi against the Trial Chamber's decision on the legality of the transfer of call data records, STL-11-01/T/AC/AR126.9/F0007-AR126.9/20150728/R001136-R001176/EN/dm, 28 July 2015 (*Prosecutor v*

Use in line with collection purpose

To be proportionate, the uses to which audio data is put should correspond to the purpose behind the collection.[94] For example, for conducting accountability work in conflict-affected regions, the use of the collected data may be limited to promoting the investigation and prosecution of international crimes. A strong justification would be needed to change the use of the data to a different purpose.

Data minimisation

The amount of personal data being processed should not be excessive when weighed against the purpose for which it was collected. Even in the context of a criminal investigation, the collection of data should not be boundless without justification.[95] Data will go through several stages of processing: collection, assessment, preservation, storage, transfer (to name a few). For each of these different stages, it is important to ask whether retaining the data for the aims being pursued is necessary.

Limits to access to the data

If only a limited category of individuals with a particular interest in the data are given access, the collection and processing of the data is more likely to be proportionate than if many individuals/organisations have access to it.[96] This is particularly so if there is not a clear reason or justification for why such individuals/organisations are given access.[97] Accordingly, when data is transferred to third parties, an assessment should be carried out to determine whether they can offer safeguards against abuse and disproportionate interference and whether their storage is secure.[98]

Collectors may wish to share with third parties the relevant information derived from the audio data in the form of a report or analysis. Doing so in a manner that excludes any individuals' personal data would avoid interfering with any individuals' privacy rights, and thereby avoid the need to ensure proportionality in the sharing process. Conversely, if the report or analysis includes information through which an individual is identified or identifiable, then a careful assessment should be carried out.

Linking datasets

Investigative work is characterised by the bringing together of information to build a picture of events. Audio data can form one piece of a bigger puzzle made up of witness testimony, open-source information (including photos and videos), call data records, satellite images, and other documentary evidence.

The linking of data is relevant for the right to privacy because the more data is brought together, the more detailed and granular the picture of a person's private sphere becomes. Non-sensitive data may become sensitive if it is cross referenced with other data and sensitive information is revealed as a result. Or,

---

Ayyash et al, Decision on Transfer of Call Data Records) para. 56. *See also* UNSC CTCED, Military Evidence Guidelines, page 18, guideline 12 which stipulates that States should 'have in place a legal and policy framework that addresses the purpose of the collection, use and storage of the information, which competent authorities may store and control data, the procedures for storing and using data, as well as existing controls and guarantees against abuses'.

[94] ECtHR, Guide to the Case-Law of the European Court of Human Rights: Data Protection (2022), paras. 117-120; *Digital Rights Ireland*, Judgment, para. 61. *See also* GDPR, Recital 50.

[95] ECtHR, Guide to the Case-Law of the European Court of Human Rights: Data Protection (2022), paras. 106-108; *Digital Rights Ireland*, Judgment, paras. 57-59.

[96] *Prosecutor v Ayyash et al*, Decision on Transfer of Call Data Records, para. 57.

[97] ECtHR, Guide to the Case-Law of the European Court of Human Rights: Data Protection (2022), paras. 223-225; *Digital Rights Ireland*, Judgment, para 62.

[98] *Big Brother Watch and Others v. United Kingdom*, Judgment, para. 362.

already sensitive data may become more so.[99] In cases where some (or all) of the data was shared voluntarily, linking it together can involve using it for a different purpose than the person intended or could have foreseen,[100] and may reveal personal characteristics the person did not intend to disclose.[101]

Because of the risks that linking data can present,[102] Collectors should be aware that additional safeguards may be needed to establish the proportionality of the data processing when data sets are cross-referenced.

4.2.3. *The Right to Privacy and Audio Data Collected by Third-Party Sources*

Collectors may come into possession of audio data because it is passed to them by third-party sources, rather than through direct collection. In this scenario, the Collectors were not involved in any potential privacy infringements that may have taken place during the initial collection and processing of the audio data. This is not to say, therefore, that Collectors can be unconcerned with this initial collection and processing; Collectors should conduct a preliminary assessment to identify any significant right to privacy issues. This involves inquiring into whether any manifest problems exist with respect to the 'in accordance with law', 'necessity', and 'proportionality' requirements discussed previously.

There are two reasons why Collectors should conduct a preliminary assessment of the actions of the third-party source: 1) because significant privacy violations can affect the admissibility of audio data as evidence before a court or tribunal, and 2) because it is relevant to the proportionality assessment that Collectors should undertake when they further process the data.

Section 5.1 below discusses in more detail when evidence will be inadmissible because it was collected in violation of human rights. Here it suffices to point out that Collectors should aim to identify possible grounds for exclusion of evidence as early as possible and flag this information to any actors they subsequently share the audio data with. Transparency is important for building robust evidence and can prevent problems in the future.

When audio data shared with Collectors by third-party sources is further processed by the Collector–meaning it is stored, analysed, enhanced, and/or transferred to other parties–an assessment of necessity and proportionality must be still undertaken. This is because the actions of the Collector in relation to the data, even if they did not collect it themselves, can infringe the right to privacy. This means that the Collector must conduct two assessments: 1) the above-stated preliminary assessment of the actions of the third-party source who directly collected the data, to check for manifest privacy issues, and 2) a detailed assessment of the Collector's own actions and plans in relation to the data. The actions of the third-party source who collected the data may be relevant for the necessity and proportionality assessment that the Collector undertakes concerning their own processing of the data. For example, if audio data was

[99] Paul Quin and Gianclaudio Malgieri, 'The Difficulty of Defining Sensitive Data–The Concept of Sensitive Data in the EU Data Protection Framework' (2021) 22 German Law Journal 1538, pages 1596-7. For a discussion on how computational capability affects data sensitivity, see Gianclaudio Malgieri and Giovanni Comandé, 'Sensitive-By-Distance: Quasi-Health Data in the Algorithmic Era' (2017) 26(3) Information, Communication and Technology Law 229.
[100] European Data Protection Supervisor, Formal consultation on EASO's social media monitoring reports (case 2018-1083), page 3.
[101] European Data Protection Supervisor, Opinion 3/2018 on online manipulation and personal data (2018), page 15.
[102] Agencia Española Protección de Datos, Risk Management and Impact Assessment in the Processing of Personal Data (2021), page 91.

collected in violation of privacy, it may make it harder to argue that the Collector's data processing is proportionate.

4.2.4. *Data Protection Instruments and Provisions*

A. Convention 108+

When personal data is involved, the Council of Europe's modernised Convention for the Protection of Individuals with Regard to the Processing of Personal Data (Convention 108+) enumerates a number of 'data subject rights' that are not fully reflected in the ECtHR case law concerning Article 8 ECHR. Generally speaking, data subject rights are the rights that a person has with regard to their personal data when such data has been processed by a third party. These rights include the right to know what personal data is being processed, the reason for the processing, and the right to object to the processing.[103] Exercising these rights requires that the individual (i.e., the data subject) knows that their personal data has been processed.

In the context of Collectors' work, data subjects (whether military or civilian) will generally not know that their data is being processed. For example, it is likely to be neither possible nor desirable for the individual whose voice is captured in an intercepted radio communication to be contacted by the Collector. Rather, Collectors can rely on Article 11 of Convention 108+, which sets out the exceptions to the protection of the rights of data subjects. These conditions mirror those that need to be satisfied for the restriction of privacy under the ECHR, so the analysis is the same here as above in section 4.2.2. The modernised Convention 108+ is also aligned with the EU's General Data Protection Regulation (GDPR), discussed in the next section.

B. The EU's General Data Protection Regulation

The EU's General Data Protection Regulation (GDPR) harmonised data privacy laws across the EU. It sets out seven general principles for data processing. These are lawfulness, fairness, and transparency; purpose limitation; data minimisation; accuracy; storage limitation; integrity and confidentiality; and accountability.[104] Many of these principles will be familiar from the discussion above in section 4.2.2.C on proportionality.

Scope of the GDPR

The territorial scope of the GDPR is broad. If the Collector has an establishment in the EU, the GDPR applies to its activities regardless of (i) the location of the data subjects (i.e., even if they are outside of EU territory); (ii) the location where the data is processed; and (iii) whether the individuals whose data is collected are EU citizens.[105] If the Collector does not have an establishment in the EU, the GDPR will still apply to its activities if it processes the data of data subjects who are in the EU.[106]

[103] Convention 108+, Article 9.
[104] GDPR, Article 5. These principles are reflected in UNESCO, Guidelines for Judicial Actors on Privacy and Data Protection (2022), page 18.
[105] GDPR, Article 3(1).
[106] GDPR, Article 3(2).

Legal Bases for Data Processing

Article 6 GDPR sets out the legal bases for data collection. Consent—where the data subject has agreed to the processing of their data—is a prominent legal basis for data collection, but will be problematic for accountability work as most data subjects are not aware of the data collection. More relevant for Collectors is the 'legitimate interests' legal basis.[107]

Data can be collected and processed if 'it is necessary for the purposes of the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data'.[108] The Information Commissioner's Office sets out a three-part test for weighing the 'legitimate interests' of the data controller (i.e., the Collector) against the 'fundamental rights and freedoms' of the data subject. The three parts are:[109]

1. Purpose test – is there a legitimate interest behind the processing? It is likely that processing data for the purposes of criminal accountability would qualify as a legitimate interest.[110]
2. Necessity test – is the processing necessary for that purpose? It must be shown that a less invasive method to achieve the legitimate interest was not possible.
3. Balancing test – is the legitimate interest overridden by the individual's interests, rights, or freedoms? This will require an assessment of the risk of harm to the individual compared with the importance of the legitimate interest.[111]

If the nature of the data processing changes—for example, the initial processing was collection and preservation, but this is later expanded to include transfer of data to third parties—then this three-part test must be newly carried out and satisfied.

Notification Requirements

When personal data is collected, the data controller[112] must inform the data subject of the identity of the controller, their contact details, the legal basis and purpose of the processing, and who the data may be transferred to, among other things.[113] Depending on whether the data was collected from the data subject themselves or not, this must be done at the point of data collection or at a later time.

The GDPR notification requirements can be onerous. Collection methods such as data scraping can collect the personal data of millions of people, in which case those millions of people must be notified of the information in the previous paragraph. Data scraping is covered by Article 14 GDPR, as this provision sets out the notification requirements for situations where personal data is not collected directly from the data subject. Article 14(5)(b) contains three limited exemptions to the notification obligation:

1) Where notifying the data subject would be impossible;
2) Where notifying the data subject would involve a disproportionate effort; or
3) Where notifying the data subject would make the achievement of the objectives of the data processing impossible or seriously impair them.

---

[107] GDPR, Article 6(1)(f).
[108] GDPR, Article 6(1)(f).
[109] 'What is the 'legitimate interests' basis?' (Information Commissioner's Office).
[110] GDPR, Recital 50.
[111] 'What is the 'legitimate interests' basis?' (Information Commissioner's Office).
[112] GDPR, Article 4(7): 'controller' means the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data.
[113] GDPR, Articles 13(1) and 14(1).

The burden of proving impossibility lies with the Collector, who must show that there are 'factors that actually *prevent it* from providing the information in question to data subjects'.[114] In order to show that notification would involve disproportionate effort, the Collector must balance the effort involved in providing the information to the data subject with the impact and effects on the data subject if they are not provided with the information. This balancing exercise should be documented and further procedural steps followed.[115] In practice, the disproportionate effort exemption has been interpreted narrowly. In a case involving a Swedish company operating in Poland, the Polish Data Protection Authority (DPA) and a Polish court determined that the financial cost of posting letters or sending text messages to nearly six million people–which the company estimated to be €8M[116]–was not sufficient to exempt it from the notification obligation.[117] The Polish court, upholding the DPA's decision, clarified that 'disproportionate effort' describes a situation where it is objectively possible but extremely difficult to notify the data subjects; financial reasons do not qualify. The act of posting a notice on the company website was also not sufficient, as the data subjects would not have known to look there given that they did not know their data had been collected.[118] It is unclear how this would play out in a criminal accountability context.

The final Article 14(5)(b) exemption is perhaps the most promising for accountability work. If a potential perpetrator must be contacted and notified that their personal data has been collected, this would make it impossible to achieve the objectives of a collection effort or, indeed, seriously impair them. This is especially so since the information that must be provided in the notification includes the identity of the Collector and the legal basis and purpose of the processing.[119] It is worth noting that these exemptions only apply when personal data is not collected directly from the data subject–Article 13 GDPR, which addresses situations where personal data is collected directly from data subjects, contains no exemptions.

Data Protection Measures

The GDPR contains a range of measures that data controllers and processors[120] must take when carrying out data processing activities. Collectors should familiarise themselves with the measures that apply to them. Some examples include:

- Implementing appropriate technical and organisational measures designed to implement data protection principles;[121]
- Implementing appropriate technical and organisational measures (for example, pseudonymisation, encryption, anonymisation, and other safeguards[122]) to ensure that only personal data that is necessary for the purpose of the processing is processed;[123]
- Keeping detailed records of the processing of data, including the names of the controllers and processors, the purpose of the processing, any transfers of data, etc.;[124]

[114] Article 29 Data Protection Working Party, Guidelines on transparency under Regulation 2016/679 (2017) (WP 29 Transparency Guidelines), page 29.
[115] These steps are detailed in WP 29 Transparency Guidelines, page 31.
[116] Natasha Lomas, 'Covert data-scraping on watch as EU DPA lays down "radical" GDPR red-line' (Techcrunch, 30 March 2019).
[117] Joanna Karolina Tomaszewska, 'Polish court overturns DPS's first GDPR fine' (International Association of Privacy Professionals, 2 April 2020).
[118] Joanna Karolina Tomaszewska, 'Polish court overturns DPS's first GDPR fine' (International Association of Privacy Professionals, 2 April 2020)..
[119] GDPR, Article 14(1).
[120] GDPR, Article 4(8): 'processor' means a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller.
[121] GDPR, Article 25(1) and (2).
[122] GDPR: Privacy by Design (Intersoft Consulting).
[123] GDPR, Article 25(2).
[124] GDPR, Article 30.

- Taking security measures to protect the data that are proportionate to the level of security risk;[125] and
- Notifying the relevant authorities of any data breaches if they occur.[126]

Data Transfers

In relation to the transfer of personal data from EU to non-EU jurisdictions, Collectors should ensure that the country to which they are transferring offers an equal level of protection. Some countries have been pre-approved as safe by the European Commission, and no further approval needs to be sought for transfer to those countries. At the time of this Protocol's publication there are 11 countries on the pre-approved list:[127] Andorra, Argentina, Canada,[128] Faroe Islands, Guernsey, Israel, Isle of Man, Japan, Jersey, New Zealand, Republic of Korea, Switzerland, the United Kingdom, the United States,[129] and Uruguay.

Where a country is not on the pre-approved list, the GDPR prohibits transfers of personal data if adequate safeguards are not put in place.[130] Such safeguards include model contract clauses (Standard Contractual Clauses, or SCCs), pre-approved by the European Commission, to be used in agreements to govern the transfer of data from controllers or processors in the EU to those outside. Additionally, companies must also ensure that the SCCs will be adhered to within the legal framework of the recipient country, or else implement additional protective measures.[131] This can be achieved via the performance of an impact assessment to analyse the risks involved in transferring the data, with consideration for the legal context of the recipient country.[132]

C. Ljubljana-The Hague Convention

The Ljubljana-The Hague Convention on International Cooperation in the Investigation and Prosecution of the Crime of Genocide, Crimes Against Humanity, War Crimes, and Other International Crimes contains a specific provision on the use and protection of personal data. This convention applies to cooperation between States and does not grant individual rights, yet is worth Collectors' consideration because it signals the importance of the right to privacy and data protection in the accountability space. The text of Article 16 of the convention reflects the GDPR principles of purpose limitation, accuracy, storage limitation, integrity and confidentiality, as well as the data subject rights to access, rectification, erasure, and notification.[133]

---

[125] GDPR, Article 32.
[126] GDPR, Article 33.
[127] European Commission, Data Protection Adequacy for Non-EU Countries
[128] Only commercial organisations.
[129] Only commercial organisations participating in the EU-US Data Privacy Framework.
[130] GDPR, Article 46 (1) and 46 (2)(c).
[131] *Data Protection Commissioner v. Facebook Ireland LTD, Maximillian Schrems*, Judgment, CJEU, C-311/18, 16 July 2020, para. 134.
[132] Guidance on the performance of such an assessment can be found in European Data Protection Board, Recommendations 01/2020 on measures that supplement transfer tools to ensure compliance with the EU level of protection of personal data (2021); *see also* Amancaya Schmitt, International Data Transfer of HR Data From the EU to Non-EU Entities – The Deadline for Adapting SCCs is December 27, 2022, (Littler, 4 October 2022).
[133] Ljubljana-The Hague Convention, Article 16.

D. African Union Convention on Cyber Security and Personal Data Protection

The AU Convention on Cyber Security and Personal Data Protection–known as the Malabo Convention–came into force in June 2023. The Convention is a framework treaty which, among other things, requires member States to implement data protection rules in their domestic legal systems. The terms of the Malabo Convention mirror, in large part, the standards in the GDPR and Convention 108+. The six basic principles of the Malabo Convention are consent, lawfulness and fairness, purpose limitation, accuracy, transparency, and confidentiality.[134] The Convention also identifies specific principles for processing sensitive data[135] and the interconnection of personal data files.[136]

4.3. The Right to Fair Trial

Individuals facing criminal charges are entitled to rights and protections under international human rights law, including the right to a fair trial.[137] This means that certain protections must be provided, including to be prosecuted within a reasonable time, by an impartial and independent judge, to have adequate time and resources to prepare a defence, and to be presumed innocent. The Collectors addressed by the Protocol are not State entities and do not themselves have the power to prosecute individuals. However, given that the audio data Collectors gather may be used as evidence in future criminal trials, the right to fair trial is relevant to their work because it can impact the evidentiary value of the data.[138] As such, while they are not legally obliged to respect the protections involved in the right to a fair trial, it is both legally and ethically desirable that Collectors incorporate consideration for these protections into their workflow.

Below is an overview of the key fair trial protections that are relevant for Collectors in their audio data work. There is an explanation of what each protection entails in general, followed by an indication of the ways in which the Collector can act accordingly.

The right to be presumed innocent	In General
	The presumption of innocence is the legal principle according to which any person accused or suspected of a crime is considered innocent until they are proven guilty.[139] The burden is on the prosecutor to prove the guilt of the accused beyond a reasonable doubt. The presumption of innocence applies from the moment an individual is identified as a suspect. Individuals do not need to have been formally charged with a crime to enjoy the right to be presumed innocent.[140] When an individual is formally charged and classified as an accused, the presumption of

[134] Malabo Convention, Article 13.
[135] Malabo Convention, Article 14.
[136] Malabo Convention, Article 15.
[137] ECHR, Article 6; ICCPR, Article 14; ACHR, Article 8; ACHPR, Article 7.
[138] See *infra* section 5.1. on 'Admissibility of Evidence' and *supra* 'Methodology: How was it Developed?'
[139] Rome Statute, Article 66; Statute of the International Residual Mechanism for Criminal Tribunals (IRMCT), Article 19 (3); UDHR, Article 11(1); ICCPR, Article 14 (2); ECHR, Article 6(2); ACHR, Article 8(2), ACHPR, Article 7(1)(b).
[140] Rome Statute, Article 66 (1) clarifies that the presumption of innocence extends to persons beyond the accused, stating that 'everyone' should be treated accordingly. See further Christoph JM Safferling, *Towards an International Criminal Procedure* (OUP 2003) page 67; Salvatore Zappala, *Human Rights in International Criminal Proceedings* (OUP 2003), page 84; and Karin N Calvo-Goller, *The Trial Proceedings of the International Criminal Court. ICTY and ICTR precedents*, (Martinus Nijhoff Publishers 2006), page 56.

	innocence continues to apply. It is only once a final decision is reached on the guilt of an accused that the presumption of innocence ceases to apply (and does not apply at the sentencing stage[141]).
	For Collectors
	• Abstain from making public statements about an individual's guilt based on the audio data collected.[142] This also applies to audio recordings containing potentially incriminating statements, as the presumption of innocence applies even if an individual admits guilt.[143] • Consider carefully who audio data is transferred to, as third parties may make prejudicial statements based on the data. This includes the media, which could publish material prejudging the individual's trial.[144] Some ways to share data with caution include redaction or anonymisation of certain information, as well an agreement with the third-party recipient to ensure the appropriate use of the audio data. • If Collector team members are called as witnesses, they should be conscious of presenting their testimony in an objective manner without making statements relating to guilt or innocence.
To know the case against them and be given exculpatory material	In General
	The accused is entitled to be informed of the evidence the prosecution intends to rely on.[145] The prosecution is obligated to disclose any evidence which may mitigate the guilt of the accused or affect the credibility of prosecution evidence.[146]
	For Collectors
	Collectors can protect this right by sharing all relevant material with either the prosecution or defence, as requested. If a request is made, a Collector should continue to disclose material it identifies as being potentially exculpatory even after the initial data handover.
To examine witnesses against them	In General
	An accused has the right to examine witnesses against them and to call witnesses in their own defence.[147]
	For Collectors

[141] *Bikas v. Germany,* Judgment, ECtHR, 76607/13, 25 January 2018, para. 57.
[142] ECtHR, Guide to the Case-Law of the European Court of Human Rights: Data Protection (2022), para. 355.
[143] *Axel Springer SE and RTL Television GmbH v. Germany*, Judgment, ECtHR, no. 51405/12, 21 September 2017, para. 51.
[144] EU Agency for Fundamental Rights, Presumption of Innocence and Related Rights: Professional Perspectives (2021), page 42; *See also*, UN Human Rights Committee, General Comment No. 32, 9-27 July 2007, which states: 'The media should avoid news coverage undermining the presumption of innocence.'
[145] Rome Statute, Article 67(2); ICC RPE, Rules 76-84; IRMCT RPE, Rules 71-73.
[146] Rome Statute, Article 67(2).
[147] Rome Statute, Article 67(1)(e); IRMCT Statute, Article 19(4)(e); ICCPR, Article 14(3)(e); ECHR, Article 6(3)(d); ACHR, Article 8(2)(f).

	Collectors can protect this right by ensuring that, when called upon, team members involved in the collection, processing, and/or analysis of audio data are available to testify for either the prosecution or defence (or both) as relevant.

A Note on the Interaction Between the Rights to Privacy and Fair Trial

Where evidence was collected and processed in a manner that violated the right to privacy, it does not automatically mean that using that evidence in a criminal trial would be unfair to the accused. If the evidence is of good quality and was collected in circumstances that do not cast doubts on its reliability and accuracy, it may potentially still be used.[148] If the following additional criteria are met, it is likely the evidence can be used without violating an accused's right to a fair trial: a) the defence is given an opportunity to challenge the authenticity of the evidence and oppose its use; b) that where the evidence is decisive in the proceedings, it is particularly strong and reliable; and c) overall, the accused's defence rights are not disregarded.[149]

5. Key Evidentiary Concepts in International Criminal Law

Audio data collected and processed by Collectors is 'information'; information can become 'evidence' if used to establish facts in legal proceedings.[150] The use of audio data as evidence has been seen in international criminal practice.[151]

Given the possibility of audio data being used as evidence in the future, this section identifies key evidentiary concepts that Collectors should keep in mind when collecting, handling, and preserving audio data. These concepts derive from the rules of evidence and case law of international criminal courts and tribunals international criminal courts and tribunals[152] and have a broad applicability and recognition. To the extent that Collectors can incorporate these concepts into their work, it will improve the evidentiary value of their audio data.

Evidence is categorised differently depending on its nature. Audio data is considered a type of documentary evidence—a broad category that comprises 'anything in which information of any description is recorded'.[153] Some forms of audio data will be contemporaneous information, meaning that it was produced at the time that the events to be proven took place. Other forms of audio data will have been created after the relevant events. Examples of contemporaneous audio data are intercepted radio and phone communications; examples of non-contemporaneous audio data are recordings of witness interviews made after the events.

---

[148] *Vukota-Bojić v Switzerland*, Judgment, ECtHR, 61838/10, 18 October 2016 (*Vukota-Bojić v Switzerland*, Judgment), paras. 94-5; *Khan v The United Kingdom*, Judgment, ECtHR, 35394/97, 12 May 2000 (*Khan v The United Kingdom*, Judgment), paras. 35-40; *Schenk v Switzerland*, Judgment, ECtHR, 10862/84, 12 July 1988 (*Schenk v Switzerland*, Judgment), paras. 47-8.
[149] *Vukota-Bojić v Switzerland*, Judgment, paras. 94-5; *Khan v The United Kingdom*, Judgment, paras. 35-40; *Schenk v Switzerland*, Judgment, paras. 47-8.
[150] OHCHR, Berkeley Protocol, page 7.
[151] *See*, most recently, *Prosecutor v Ongwen*, Appeal Judgment, ICC-02/04-01/15-2022-Red, 15 December 2022 (*Prosecutor v Ongwen*, Appeal Judgment) and *Prosecutor v Al-Hassan,* Judgment, ICC-01/12-01/18-2594-Red, 26 June 2024 (*Prosecutor v Al-Hassan,* Judgment).
[152] When working with a particular domestic jurisdiction, there may be additional and possibly more stringent evidentiary standards to consider and comply with.
[153] *Prosecutor v. Musema*, Judgment and Sentence, ICTR-96-13-T, 27 January 2000, para. 53.

This section will first cover admissibility of evidence, clarifying what factors may (or may not) impact a court's decision to admit an item of evidence at trial. It will then address the evidentiary concepts of relevance, probative value, reliability, authenticity, prejudicial effect, and weight of an item of evidence, assessing which elements make a piece of evidence more persuasive in proving an issue at trial. The section closes with information on the possibility of Collector team members being called as witnesses.

5.1. Admissibility of Evidence

For audio data to be relied upon as evidence in a criminal trial, it must first be admitted into evidence pursuant to the rules of procedure and evidence of the court. The paragraphs below set out core considerations relating to the admission of audio data as evidence.

A three-part test is used to determine the admissibility of an item of documentary evidence.[154] The court must determine that:

- The evidence is *prima facie* relevant (see section 5.2.);
- The evidence has *prima facie* probative value (see section 5.3.); and
- Where relevant, the prejudicial effect of the evidence does not outweigh its probative value (see section 5.4.).[155]

The admissibility of an item of evidence has no bearing on the final weight to be afforded to it (see section 5.5.).[156] Generally, the international standards for admissibility are permissive[157] and evidence is unlikely to be excluded.[158]

Despite this permissive standard, international criminal law does contain some exclusionary rules. The Rome Statute of the ICC establishes that evidence will not be admissible if

A. it has been obtained in violation of the Statute or 'internationally recognized human rights'; and
B. the violation 'casts substantial doubt on the reliability of the evidence' (the concept of reliability is dealt with in section 5.3.), or the admission of the evidence 'would be antithetical to and would seriously damage the integrity of the proceedings'.[159]

---

[154] *See*, for example, *Prosecutor v Lubanga*, Corrigendum to Decision on the admissibility of four documents, ICC-01/04-01/06, 20 January 2011, paras. 28–31; *Prosecutor v Bemba*, Public redacted version of the First decision on the prosecution and defence requests for the admission of evidence, dated 15 December 2011, ICC-01/05-01/08-2012-Red, 9 February 2012, para. 13.
[155] *See*, for example, *Prosecutor v Lubanga*, Corrigendum to Decision on the admissibility of four documents, ICC-01/04-01/06, 20 January 2011, paras. 28–31; *Prosecutor v Bemba*, Public redacted version of the First decision on the prosecution and defence requests for the admission of evidence, dated 15 December 2011, ICC-01/05-01/08-2012-Red, 9 February 2012, para. 13.
[156] *Prosecutor v Bemba*, Decision on the admission into evidence of items deferred in the Chamber's "Decision on the Prosecution's Application for Admission of Materials into Evidence Pursuant to Article 64(9) of the Rome Statute" (ICC-01/05-01/08-2299), ICC-01/05-01/08, 27 June 2013 (*Prosecutor v Bemba*, Decision on the admission into evidence of deferred items), para. 9.
[157] The ICTY and ICTR legal frameworks establish that chambers 'may admit any relevant evidence which it deems to have probative value' (ICTY RPE, Rule 89(C); ICTR RPE, Rule 89(C)). The ICC framework provides that judges freely assess all types of evidence submitted, enjoying in this respect 'a significant degree of discretion' (ICC RPE, Rule 63(2)); *See*, *inter alia*, *Prosecutor v Lubanga,* Decision on the admissibility of four documents, ICC-01/04-01/06, 13 June 2008, para. 24.
[158] See Diletta Marchesi, 'Intercepted Communications in the Ongwen Case: Lessons to Learn on Documentary Evidence at the ICC' (2021) 22 International Criminal Law Review 1 page 8 ('the admissibility standard for documentary evidence is not only 'permissive', but almost absent.');Yvonne McDermott, Proving International Crimes (OUP 2024) page 64 ('Trial Chambers have taken an inclusive approach to the admissibility of evidence, with any weaknesses in the evidence more likely to speak to its ultimate weight than to its probative value for the purposes of admissibility'); Nikita Mehandru and Alexa Koenig, 'Open Source Evidence and the International Criminal Court' (2019) Harvard Human Rights Journal ('While information will often be admitted as evidence if shown to be even remotely relevant, the weight that judges will accord that information may vary.')
[159] Rome Statute, Article 69(7).

A and B are cumulative requirements, and the threshold for B is relatively high. For example, the failure of domestic authorities to comply with domestic procedural law when collecting evidence would not, in and of itself, be enough to damage the integrity of ICC proceedings (even in a situation where this breach of domestic law resulted in a violation of the right to privacy).[160] A classic example of evidence being both unreliable and antithetical to the integrity of proceedings, and thus meeting the said threshold, would be a confession obtained through torture.

Other international criminal courts and tribunals international criminal courts and tribunals have exclusionary rules similar to those of the ICC,[161] and the threshold for exclusion has been similarly high. When it comes to phone intercepts, case law from other courts and tribunals shows that intercepts obtained illegally per domestic law will not necessarily be inadmissible.[162] Rather, case law suggests that the manner and circumstances in which they were obtained, as well as their reliability and effect on the integrity of the proceedings, will be factored into the assessment of their admissibility.[163] The case law has further underscored that, in situations of armed conflict in particular, intelligence that may result from illegal activity may prove essential in uncovering the truth, especially when this information is not available from other sources.[164] Intercepts have been generally admitted unless the integrity of the proceedings would be seriously damaged.[165]

With regard specifically to data that has been hacked and then leaked, there is no rule that such data would be *per se* inadmissible. Hacked data is data 'acquired by an outsider who gains unauthorised access to it', and leaked data is understood as data 'obtained by an insider who has authorised access to it, but shares it in an unauthorised manner'.[166] Hacked or leaked data may be inadmissible if its authenticity cannot be established[167] or if the above exclusionary rule applies.[168]

---

[160] *See Prosecutor v Bemba et al*, Decision on Request in Response to Two Austrian Decisions, ICC-01/05-01/13-1948, 14 July 2016. *See also Prosecutor v Bemba et al,* Decision on Requests to Exclude Western Union Documents and other Evidence Pursuant to Article 69(7), ICC-01/05-01/13-1854, 29 April 2016 and *Prosecutor v Bemba et al,* Appeal Judgment, para. 289. *See further* Petra Viebig, *Illicitly Obtained Evidence at the International Criminal Court* (TMC Asser Press, 2016) pages 147-149.

[161] ICTY RPE, Rule 96; ICTR RPE, Rule 95; STL RPE, Article 162; RMICT RPE, Rule 117; Kosovo Specialist Chambers (KSC) RPE, Rule 138.

[162] In *Prosecutor v Renzaho,* Decision on Exclusion of Testimony and Admission of Exhibit, paras. 15–16, the ICTR held that the tape of a call between Rwandan authorities which was intercepted by Rwandan Patriotic Front soldiers using a walkie-talkie and simultaneously recorded by a journalist was not 'antithetical to and certainly would not seriously damage the integrity of the proceedings'. In *Prosecutor v Brdjanin*, Decision on the Defence "Objection to Intercept Evidence", paras. 53-5, the ICTY underlined that its jurisprudence had never upheld the exclusionary rule as a matter of principle: the right to privacy can be derogated from in times of emergency as in the course of a war, hence 'communications intercepted during an armed conflict are not as such subject to exclusion under Rule 95 and should therefore be admitted'. In *Prosecutor v Mbarushimana*, Decision on the confirmation of charges, ICC-01/04-01/10-465-Red, 16 December 2011, para. 71, the ICC rejected a Defence challenge to the admissibility of intercepted phone calls, in part, because the Defence had failed to make 'any submissions to the effect that a lack of authorisation of the intercepts would have any impact on the reliability of the evidence thereby obtained, or that their admission into evidence would be antithetical to or would seriously damage the integrity of the proceedings'.

[163] *Prosecutor v Brdjanin*, Decision on the Defence "Objection to Intercept Evidence", para. 55.

[164] *Prosecutor v Brdjanin*, Decision on the Defence "Objection to Intercept Evidence", para. 61; *Prosecutor v. Kordić and Čerkez*, Public Transcript of Hearing 2 February 2000, ICTY Case No. IT-95-14/2, 2 February 2000, page 13694. See also M. Klamberg, *Evidence in International Criminal Trials, Confronting Legal Gaps and the Reconstruction of Disputed Events* (Martinus Nijhoff, 2013), pages 395-406.

[165] *Prosecutor v Brdjanin*, Decision on the Defence "Objection to Intercept Evidence", paras. 61 and 63.

[166] Lindsay Freeman, 'Hacked and Leaked: Legal Issues Arising From the Use of Unlawfully Obtained Digital Evidence in International Criminal Cases' (2021) 25(2) UCLA Journal of International Law and Foreign Affairs 45 (Freeman, Hacked and Leaked), page 47.

[167] As was the case in *The Prosecutor v. Ayyash, Badreddine, Merhi, Oneissi, and Sabra,* Decision on the Admissibility of Documents Published on the Wikileaks Website, STL-11-01/T/TC/F1955/20150521/R274176-R274189/EN/af, 21 May 2015, paras 40-43.

[168] See Rome Statute, Article 69(7). For a detailed discussion on when hacked and leaked data may be inadmissible on these (and other) grounds, *see generally* Freeman, Hacked and Leaked.

In addition to sharing audio data with international courts and tribunals, Collectors may intend to transfer collected audio data to domestic authorities. In the latter case, Collectors should examine the admissibility requirements applicable in that particular jurisdiction.

5.2. Relevance of Evidence

An item of evidence is relevant when it pertains to the matters considered at trial. In other words, when it can be used to show that the existence of a particular fact is more or less probable. If evidence is irrelevant, it can be ruled inadmissible; or, if it has low relevance, it will likely be ascribed less weight by the judges.

Audio data will likely be considered relevant by international criminal courts and tribunals if it constitutes a contemporaneous, chronological record of events on the ground relevant to the charges.[169] If a member of the Collector's team involved in the collection process can testify to the fact that the recordings are a contemporaneous record (through her or his personal recollection, the metadata, and other documentation, etc.), this will contribute positively to a finding that the audio data is relevant to the case at hand.[170] Accurate documentation of the date on which the audio data was collected is key to the data's relevance, as this is necessary to show that the data relates to the period of time the criminal charges are concerned with.[171] Likewise, information as to the location where the audio data was recorded will be important for relevance.[172]

When it comes to determining relevance, intelligibility is fundamental. This means that the relevant material should be presented in an understandable format (readable and/or audible). For example, raw or non-demodulated radio signals would be considered unintelligible and, thus, may be deemed inadmissible.[173] Moreover, the material can be reviewed for relevance only if it is presented in one of the working languages of the relevant court, meaning that translations are fundamental and translated transcripts will likely be required.[174] Collectors should bear in mind that the use of professional translators may be needed to support automated translations. Where only the ambient sound of the audio data is relevant (and not human voices) it is possible that transcription will not be necessary.[175]

Logbooks (in a working language of the relevant court) have also been reviewed by courts and considered 'an essential part of the [...] assessment of particular recordings' as they were deemed 'contemporaneous

---

[169] *Prosecutor v Mladić*, Decision on Prosecution's Bar Table Motion for the Admission of Intercepts: Srebrenica Segment, IT-09-92, 2 May 2013, (*Prosecutor v Mladić,* Decision on Admission of Intercepts), para. 24.

[170] *Prosecutor v Popović et al.,* Judgment Volume I, ICTY, Case No. IT-05-88-T, 10 June 2010 (*Prosecutor v Popović,* Judgment), para. 65.

[171] *Prosecutor v Ntaganda,* Decision on Prosecution's request for admission of documentary evidence, ICC-01/04-02/06-1838, 28 March 2017, para. 68; *Prosecutor v. Bemba*, Public Redacted Version of "Decision on the Prosecution's Application for Admission of Materials into Evidence Pursuant to Article 64(9) of the Rome Statute" of 6 September 2012, ICC-01/05-01/08-2299-Red, 8 October 2012 (*Prosecutor v. Bemba*, Public Redacted Version of "Decision on Admission of Materials"), para. 84; *Prosecutor v Katanga and Chui*, Decision on Bar Table Motion, para. 24.

[172] *Prosecutor v Katanga and Chui*, Decision on Bar Table Motion, para. 24.

[173] This can be inferred by analogy from the case law of international criminal courts and tribunals concerning call data records—a form of metadata that provides information about the communication, including source, date, time and duration of the call. Call data records have been considered inadmissible on the basis of their unintelligibility. See *Prosecution v Ayyash et al.,* Judgment, STL-11-01/T/TC, 18 August 2020, paras. 375–378, where the Trial Chamber rejected the admission of call data records due to it being voluminous and unreadable and containing a string of numbers and symbols. *See also*, Leiden Guidelines on the Use of Digitally Derived Evidence in International Criminal Courts and Tribunals (Leiden Guidelines), Guideline E1 on call data records.

[174] *Prosecutor v Ongwen*, Trial Judgment, ICC-02/04-01/15-1762-Red, 4 February 2021 (*Prosecutor v Ongwen,* Trial Judgment), paras. 648, 650.

[175] *Prosecutor v Mladić*, Transcript, IT-09-92 (19 September 2012), page 2634.

written records' of the intercepted communications.[176] Collectors should thus be ready to provide transcripts and translations of the audio recordings as well as their contemporaneous documentation of the recording process in an intelligible manner. While desirable, it may not always be necessary for the court to have access to the original audio recordings when enough evidence surrounding the audio data (e.g., corroborative testimonial evidence and large amounts of documentary evidence) exists.[177]

Case law demonstrates that relevance is not necessarily affected by minor time and date discrepancies[178] or by disagreements between the parties as to how to interpret what is said in an audio recording.[179] Possible discrepancies between the original and the translated versions of the audio recording are also not an obstacle to their relevance *per se*.[180] These issues contribute to the weight of the evidence, rather than the relevance. Collectors should therefore not discount audio data on these bases.

5.3. Probative Value of Evidence

An item of evidence is probative when it has the potential to prove or disprove an asserted fact. Probative value is not synonymous with relevance, although the two are sometimes confused.[181] Relevance is a 'yes or no' question, while probative value is more of a spectrum: a relevant item of evidence can have lower or higher probative value.

An assessment of the probative value of a piece of evidence is seldom based on a single factor. The following have been considered in the ICL case law as enhancing the probative value of audio data: a) where it was collected from;[182] b) the provision of the original audio recording to the court alongside transcripts and translations; c) the fact that a voice on an audio recording has been identified as the defendant; and d) the confirmation of the information in the audio recording by a witness.[183] The probative value of audio evidence can also be bolstered by additional information that can, for example, help to identify the voice that is heard in the recording, or establish the circumstances of the audio recording's creation, preservation, and handling.[184]

Advanced technological capacity and professional expertise are not required to provide probative value to audio data.[185] That said, for the data to have higher probative value, it may have to be shown that the

[176] *Prosecutor v Ongwen*, Trial Judgment, para. 658.
[177] *Prosecutor v Blagojević and Jokić*, Decision on Admission into Evidence of Intercept-Related Materials, ICTY, Case No. IT-02-60-T, 18 December 2003 (Prosecutor v Blagojević and Jokić, Decision on Admission of Intercept-Related Materials), para. 25.
[178] *Prosecutor v. Mladić*, Decision on Prosecution Motion to Admit Evidence from the Bar Table: Excerpts from Mladic's Audio Tapes, ICTY, Case No. IT-09-92-T, 18 September 2013, (*Prosecutor v Mladić,* Decision on Excerpts from Mladic's Audio Tapes), para. 8.
[179] *Prosecutor v Mladić*, Decision on Excerpts from Mladic's Audio Tapes, para. 9.
[180] *Prosecutor v. Popović et al.,* Decision on Admissibility of Intercepted Communications, ICTY, Case No. IT-05-88-T, 7 December 2007, paras. 75 and 78.
[181] *See*, e.g., *Prosecutor* v. *Bemba*, Decision Pursuant to Article 61(7)(a) and (b) of the Rome Statute on the Charges of the Prosecutor against Jean-Pierre Bemba Gombo, ICC-01/05-01/08, 15 June 2009, para. 42, where it states that an item of evidence is considered relevant if it has probative value.
[182] Their probative value was justified by the fact that they had been recovered by the Serbian authorities from the residence of the defendant's family (*Prosecutor v Mladić*, Decision on Excerpts from Mladic's Audio Tapes, para. 9).
[183] *Prosecutor v. Mladić,* Decision on Prosecution Motion for Admission of Documents from the Bar Table, paras. 11-12.
[184] Leiden Guidelines, Guideline F4, citing *Prosecutor v. Bemba*, Public Redacted Version of "Decision on Admission of Materials", paras. 84, 119, and 121.
[185] *Prosecutor v Mladić,* Judgment Volume IV of V, IT-09-92-T, 22 November 2017, paras. 5305-5307; *Prosecutor v. Mladić*, Transcript, IT-09-92, 13 August 2015, pp. 37746–37747.

person who obtained/collected it had the technical means to do so.[186] Collectors should therefore ensure that team members involved in audio data collection are adequately trained and keep up to date with the latest technological developments.

Information regarding the provenance of audio data is particularly important in determining its probative value. Provenance 'relates to the origin or earliest known existence of something'.[187] For audio data, provenance can mean different things depending on the data source. For intercepted communications, the actor making the audio recording of the communication will be the creator–or the author–of the data. Collectors engaged in interception can therefore speak to the provenance of the resulting audio data. The same is true for situations where a Collector is recording an interview with a witness or a live radio broadcast. By contrast, when Collectors are engaged in open-source data collection, for example from social media sites, or when an audio recording is sent to them over an instant messaging app, the author of the data is a third party. Additional steps will therefore be needed to establish and adequately record details of the data's provenance. For content posted to social media, information such as the uploader, the page to which material was posted, and the post title and description can help to establish provenance, which speaks to reliability and authenticity, and by extension, probative value.[188]

Being unable to provide information about audio data's provenance, or providing only limited information, will negatively impact its probative value, in some cases to the extent that the data will be inadmissible.[189] Ideally, the author of the data would testify in court with the data being submitted into evidence through them as a witness; this will improve the probative value of data by allowing for the scrutiny of cross examination.[190] While there is no blanket prohibition on admitting evidence where its purported author has not been called to testify, factors such as 'proof of authorship will naturally assume the greatest importance' in judges' determination of the weight of evidence.[191] Weight, as is detailed below in section 5.5., is the relative importance attached to an item of evidence in deciding whether a certain issue has been proven or not.[192]

Reliability of evidence as an element of probative value

Reliability is an important part of the admissibility test as it is necessary for probative value. Unreliable evidence will not be admissible. An item of evidence is considered reliable if the veracity and accuracy of its content can be trusted. Accordingly, an item of evidence is deemed insufficiently reliable if it cannot be said to prove or disprove a relevant assertion.[193]

Different factors are relevant when assessing reliability. It must be ascertained whether the evidence displays such qualities that, when considered alone, could reasonably be believed.[194] The case law has

[186] Leiden Guidelines, Guideline D5. The Trial Chamber in *Prosecutor v Blagojević and Jokić*, Decision on Admission of Intercept-Related Materials noted the experience of the interceptors, their professional certifications, and how long standing their experience in the conflict was (para. 22). The Chamber in *Prosecutor v Ongwen*, Trial Judgment also noted the qualification of the witnesses, namely that they were either the intercept operators or their supervisors (multiple paragraphs).

[187] OHCHR, Berkeley Protocol, page 63.

[188] *Prosecutor v Al-Hassan*, Judgment, para. 726 fn. 2175.

[189] *Prosecutor v Renzaho,* Decision on Exclusion of Testimony and Admission of Exhibit.

[190] *Prosecutor v Delalic*, Decision on the Motion of the Prosecution for the Admissibility of Evidence, ICTY Case No. IT-96-21, 19 January 1998 (*Prosecutor v Delalic*, Decision on Admissibility of Evidence), para 22.

[191] *Prosecutor v Delalic*, Decision on Admissibility of Evidence, paras 20-22, cited with approval in *Prosecutor v Brdjanin and Talić,* Order on the Standards Governing the Admission of Evidence, ICTY Case No. IT-99-36-T, 15 February 2002, para 18.

[192] *Prosecutor v Katanga and Chui*, Decision on Bar Table Motion, paras. 13-14.

[193] *Prosecutor v Katanga and Chui*, Decision on Bar Table Motion, para. 28.

[194] *Prosecutor v Katanga and Chui*, Decision on Bar Table Motion, para. 26.

held that 'there is no finite list of possible criteria that are to be applied in determining reliability.'[195] The following factors are important in the assessment of items of evidence:

a) the source of the information, in particular whether the source has an allegiance towards one of the parties in the case or has a personal interest in the outcome of the case, or whether there are other indicators of bias;
b) the nature and characteristics of the item of evidence (e.g., whether the evidence is an audio or video recording, the public or private character of the information, etc.);
c) the contemporaneous nature of the evidence (i.e. whether the information was obtained and recorded simultaneously or shortly after the events to which it pertains);
d) the purpose of the item (whether the document was created for the specific purpose of criminal proceedings or for some other reason); and
e) adequate means of evaluation (whether the information and the way in which it was gathered can be independently verified or tested).[196]

Consistency, clarity, and transparency in the collection process are crucial factors for the reliability of audio data.[197] Collectors must therefore ensure that their collection processes are documented in great detail[198] and applied consistently, with changes to the documentation process explained where needed. Maintaining a detailed record can help to overcome any evidentiary shortcomings that may arise throughout the collection process.[199]

Careful recordkeeping helps to establish chain of custody, which in turn is important to proving reliability. Chain of custody 'refers to the chronological documentation of the sequence of custodians of a piece of information or evidence, and documentation of the control, date and time, transfer, analysis and disposition of any such evidence'.[200] An imperfect chain of custody does not necessarily render the data inadmissible, but the chain of custody concerns will factor into the reliability and probative value assessment.[201]

Where the author of the data is the Collector—for example, where it is the Collector who has made the audio recording—demonstrating chain of custody is reasonably straightforward. It involves keeping a careful record of who had access to the data and of any time the data was accessed, moved, transferred, or in any way altered or copied. Where the Collector is not the author, efforts should be made to document the data's chain of custody before it was collected by the Collector. Where audio data has been transferred

[195] *Prosecutor v Katanga and Chui*, Decision on Bar Table Motion, para. 27.
[196] *See Prosecutor v Katanga and Chui*, Decision on Bar Table Motion, para. 27.
[197] *Prosecutor v Tolimir*, Judgment, ICTY Case No. IT-05-88/2-T, 12 December 2012 (*Prosecutor v Tolomir,* Judgment), para. 64, referring to *Prosecutor v Tolimir,* Transcript, ICTY, Case No, IT-05-88/2-T, 7 September 2010, p. 5033; *Prosecutor v Blagojević and Jokić*, Decision on Admission of Intercept Materials, para. 21. *See also Prosecutor v Katanga and Chui*, Decision on Bar Table Motion, para 30 where, in relation to reports composed by NGOs, the Chamber stressed the importance of providing information on 'their sources and the methodology used to compile and analyse the evidence upon which the factual assertions are based'. Without this information, the Chamber 'cannot assess the reliability of the content of the reports'.
[198] *Prosecutor v Ongwen*, Decision on the confirmation of charges against Dominic Ongwen, ICC-02/04-01/15-422-Red, 23 March 2016 (*Prosecutor v Ongwen*, Confirmation of Charges), para. 51; the Berkeley Protocol furthermore provides that '[i]nvestigators should document their activities during each phase. This will help with the understandability and transparency of their investigations, including chains of custody, and with the efficiency and efficacy of their investigations, including completeness and communication among team members' (OHCHR, Berkeley Protocol, page 53).
[199] *Prosecutor v Ongwen*, Confirmation of Charges, para. 51; *Prosecutor v Ongwen*, Transcript, ICC-02/04-01/15-T-20-Red-ENG, 21 January 2016, para. 44, lines 8-24.
[200] OHCHR, Berkeley Protocol, page 61.
[201] *Prosecutor v Brdjanin*, Judgement, ICTY Case No. IT-99-36-T, 1 September 2004 (Prosecutor v *Brdjanin*, Judgement), para 34. In this case, what made the chain of custody less than perfect was the fact that they were stored in the unsupervised possession of a person for more than 10 years before being transferred to the ICTY.

to the Collector by an official source, such as law enforcement authorities, securing an official document confirming chain of custody will help to establish reliability.[202] Alternatively, identifying witnesses who can speak to the provenance and chain of custody of the audio data will be helpful,[203] as will witness testimony identifying the voices on the audio recording.[204]

Detailed record keeping is also of particular importance when audio data is subject to enhancement. Enhancement aims to improve 'the quality of the original audio material to enable members of the court to comprehend or interpret the material to the best possible standard without adding to or detracting from the content of the original'.[205] The following factors have been found to improve the reliability of enhanced audio data:

a) enhancements made on a duplicate of the audio data, keeping the original preserved for comparison purposes;
b) corresponding labels on the original and the enhanced copy;
c) corresponding track times on the files;
d) broadly corresponding contents between the original and the enhanced copy when compared side by side; and
e) submission of a technical report describing the exact processes applied to each piece of data.[206]

In addition to the above, reliability is boosted when two or more team members have collected the same audio data, with only slight or no variations from each other.[207] This particularly applies when team members work from different locations,[208] or when the communication is recorded by listening devices in different locations. When multiple logbooks for a particular communication exist (for example, because the same audio data was collected from different locations), a 'word-for-word mirroring' is not expected for corroborative effect. Instead, 'differences in details are to be expected'—considering possible differences in the levels of experience of the intercepting personnel, the potential for varying quality of what could be heard at the interception site, and the fact that different people will inevitably summarise long conversations in different ways and focus on different parts.[209] Discrepancies in the logs do not necessarily require them to be corroborated by the transcripts of the audio recordings or testimony of the intercept personnel, nor do such discrepancies render the logbooks unreliable.[210]

Logs are not viewed as 'verbatim transcripts' of the audio recordings and are generally considered together with transcripts, witness testimonies, and other logbooks to verify the conversations' accuracy and the meaning.[211]

---

[202] *Prosecutor v Ngirabatware*, Decision on the Third Defence Motion for Admission of Documentary Evidence, ICTR Case No. ICTR-99-54-T, 4 July 2012, para 46. For further background on this example, see *Prosecutor v Ngirabatware*, Defence Reply to Prosecution Response to Defence Motion for Admission of Documentary Evidence, ICTR Case No. ICTR-99-54-T, 14 March 2012, paras 23-29.
[203] *Prosecutor v Popović,* Judgment*,* paras 64-65
[204] *Prosecutor v Brdjanin*, Judgment, para 34.
[205] *Prosecutor v Ongwen*, Trial Judgment, para. 651.
[206] *Prosecutor v Ongwen*, Trial Judgment, paras. 654-655.
[207] *Prosecutor v Krstić*, Judgment, ICTY, Case No. IT-98-33-T, 2 August 2001 (*Prosecutor v Krstić*, Judgment), para. 108; *Prosecutor v Blagojević and Jokić,* Decision on Admission of Intercept Materials, paras. 24 and 26.
[208] *Prosecutor v Krstić*, Judgment, para. 108; *Prosecutor v Blagojević and Jokić*, Decision on Admission of Intercept-Related Materials, paras. 24 and 26.
[209] *Prosecutor v Ongwen*, Trial Judgment, para. 664.
[210] *Prosecutor v Ongwen*, Appeal Judgment, para. 593.
[211] *Prosecutor v Ongwen*, Appeal Judgment, para 597; *Prosecutor v Ongwen*, Trial Judgment, para. 558.

Reliability can be improved by other kinds of corroborating evidence, such as the testimony of the team members who collected and worked with the audio files.[212] As noted in section 5.6., team members should be prepared to be called as witnesses to explain the Collector's processes. Non-witness evidence is also significant for improving reliability through corroboration: documents, reports, aerial images, and photographs have all played a corroborating role in the case law.[213] Corroborating evidence can help to overcome reliability challenges that arise from uncertainties in the chain of custody. For example, if a digital file's metadata indicates the data may have been accessed by a third party or external software, this does not *per se* mean that the audio itself has been modified; rather, in such a case, witness evidence and other corroborating evidence will be needed to support the data's reliability.[214]

Finally, Collectors should bear in mind that their impartiality (or lack thereof) is relevant to an assessment of reliability.[215] Without sufficient guarantees of 'non-partisanship and impartiality', information collected by NGOs may not be deemed *prima facie* reliable and might not be admissible.[216]

Authenticity of evidence as an element of reliability

Authenticity relates to whether a piece of evidence is what it professes to be in origin or authorship.[217] Authenticity is an important indicator of reliability. Several factors may be considered when assessing authenticity. Relevant considerations include whether there is verifiable information regarding the source, evidence of originality and integrity of the content,[218] and a preserved and documented chain of custody. Information about the date and the author have also been considered important.[219]

Witness testimony is crucial for establishing authenticity. The relevant Collector team members should, therefore, be prepared to be called to testify—for example, to confirm that they recognise the recording and associated transcripts and to confirm that it is the same recording as the one they were involved in collecting.[220] In addition, an individual's self-identification during the intercepted communication can serve as an inherent indicator of the communication's authenticity.[221]

Certain factors that contribute to the reliability of evidence can also contribute to its authenticity. In relation to audio data, authenticity can be enhanced by other independent corroborative evidence, such as recordings made by others of the same conversations or other documentary evidence.[222]

---

[212] *Prosecutor v Ongwen*, Confirmation of Charges, para. 51; *Prosecutor v Ongwen*, Trial Judgment, para. 643.
[213] *Prosecutor v Krstić*, Judgment, paras. 114-116; *Prosecutor v Blagojević and Jokić*, Decision on Admission of Intercept Materials, para. 24.
[214] *Prosecutor v Al-Hassan,* Judgment, para. 811 fn. 2572.
[215] The ICC's Trial Chamber II, when deciding on the admissibility of two UN reports, considered it relevant to a determination of probative value that the reports were 'established by UN services acting in an impartial manner with a concern to understand the events in question' (*Prosecutor v Katanga and Ngudjolo*, Transcript, ICC-01/04-01/07-T-229-Red-ENG, 7 December 2010, page 24).
[216] *Prosecutor v Katanga and Chui*, Decision on Bar Table Motion, para 30, cited with approval in *Prosecutor v. Bemba*, Public Redacted Version of "Decision on Admission of Materials", para 35, which in turn was cited with approval in *Prosecutor v Bemba*, Judgment pursuant to Article 74 of the Statue, ICC-01/05-01/08-3343, 21 March 2016 (*Prosecutor v. Bemba*, Judgment), para 270.
[217] *Prosecutor v Prlić et al*, Decision on Jadranko Prlić's Interlocutory Appeal against the Decision on Prlić Defence Motion for Reconsideration of the Decision on Admission of Documentary Evidence, ICTY Case No. IT-04-74-AR73.I6, 3 November 2009, para. 34.
[218] *Prosecutor v Katanga and Chui*, Decision on Bar Table Motion, para. 24.
[219] Leiden Guidelines, Guideline F4.
[220] *Prosecutor v Renzaho,* Decision on Exclusion of Testimony and Admission of Exhibit, para. 13.
[221] Leiden Guidelines, Guideline E4, citing *Prosecutor v Bemba et al.*, Judgment pursuant to Article 74 of the Statute, ICC-01/05-01/13-1989-Red, 19 October 2016, para. 219.
[222] *Prosecutor v Tolimir*, Judgment, paras. 63–66; *Prosecutor v Krstić*, Judgment, para. 108; *Prosecutor v Blagojević and Jokić*, Decision on Admission into Evidence of Intercept-Related Materials, para. 24.

5.4. Prejudicial Effect of Evidence

Under the third and last step of the admissibility test, the court must, where relevant, balance the probative value of the item of evidence under consideration against any prejudicial effect that its admission may cause to the fairness of the proceedings as a whole, and in particular, to the rights of the accused. The item will be excluded if its relevance and probative value are insufficient (or insufficiently substantiated[223]) to justify its admission in light of its potentially prejudicial effect.[224]

This balancing exercise is to be done on a case-by-case basis. For example, evidence of prior criminal or immoral conduct may suggest that the defendant is of bad character or prone to commit a crime; if the conduct in question is not strictly related to the charges, the probative value of such evidence may not justify the admission of the evidence because of its prejudice to the defendant's fair trial rights.

5.5. Weight of Evidence

Weight is a similar but distinct concept from probative value. It is the relative importance attached to an item of evidence in deciding whether a particular issue has been proven or not.[225] When determining weight, items of evidence are weighed against each other.[226] Judges usually determine the final weight to be accorded to the evidence when assessing the evidence as a whole at the end of the case.[227]

5.6. Collector Personnel Serving as Witnesses in Criminal Proceedings

It is preferable, where possible, for audio data to be submitted into evidence through a witness.[228] Operators involved in collecting audio data, their supervisors, and prosecution office staff are often called as witnesses to testify in court about the collection operations and specific communications.[229] In the ICC system, among others, witness testimony can also be introduced through prior recorded statements.[230]

Not all persons involved in data collection operations need to testify at trial.[231] Nevertheless, Collector team members should be prepared to testify as witnesses (or, if applicable and appropriate, to provide prior recorded statements) regarding their collection and processing activity. Collector personnel may be called to testify before one or more different accountability mechanisms.

Determining which individual is best suited to provide evidence will depend on various factors, including individuals' respective roles in the collection process and degrees of expertise. Testimony will likely focus less on the content of the communication collected and more on the process of collection, processing, storing, and so on to establish the reliability of the audio data.

---

[223] *Prosecutor v. Bemba*, Public Redacted Version of "Decision on Admission of Materials", para. 122.
[224] *Prosecutor v Bemba*, Decision on the admission into evidence of deferred items, para. 9.
[225] *Prosecutor v Katanga and Chui*, Decision on Bar Table Motion, paras. 13-14.
[226] Fergal Gaynor, Dov Jacobs, Mark Klamberg, and Vladimir Tochilovsky, 'Law of Evidence', in Göran Sluiter, Håkan Friman, Suzannah Linton, Sergey Vasiliev, and Salvatore Zappalà (eds), *International Criminal Procedure: Principles and Rules* (OUP 2013), p. 1027.
[227] *Prosecutor v Bemba*, Decision on the admission into evidence of deferred items, para. 9.
[228] *Prosecutor* v. *Ruto and Sang,* Decision on the Joint Defence Application for Admission of Documentary Evidence Related to the Testimony of Witness 536, ICC-01/09-01/11, 15 July 2014, para. 11.
[229] See *Prosecutor v Ongwen*, Trial Judgment, para. 555; *Prosecutor v Tolimir*, Judgment, para. 63.
[230] For instance, in the *Ongwen* trial before the ICC, the Trial Chamber allowed prior recorded testimonies from a total of 49 witnesses to be introduced (*Prosecutor v Ongwen*, Trial Judgment, para. 254). *See also*, ICC RPE, Rule 68.
[231] *Prosecutor v Ongwen,* Appeal Judgment, para. 562.

Where there are concerns for the safety and well-being of witnesses, courts can order protective measures, including

- Face/voice distortion while the witness is giving evidence;
- Use of a pseudonym;
- Conducting parts of hearings in private or closed sessions; and
- Prohibiting the Prosecution, the Defence, and any other participant in the proceedings from disclosing identifying information to a third party.

GLOSSARY

Term	Definition
Accountability mechanism	Accountability mechanisms include judicial systems, such as international or domestic courts and tribunals, whether via criminal prosecution or civil lawsuit; as well as non-judicial systems, such as truth and reconciliation commissions, treaty monitoring bodies, and ombudsmen and human rights commissions. They are designed to ensure that individuals, organisations, or institutions comply with and are held responsible for their actions and decisions. These mechanisms aim to make sure that parties' actions align with established law, standards, or commitments and that there are consequences for failing to adhere to these obligations.
Admissibility (of evidence)	The quality of being acceptable or valid as evidence in a court of law.
Algorithm	A well-defined procedure or set of instructions that allows a computer to solve a problem or respond to a predetermined scenario.[232]
Algorithmic training	A process of using a data set to train a machine learning model's parameters and omptimise its ability to analyse new data and perform requested outputs.
Anonymisation	The process of making it impossible to identify a specific individual.[233] In the context of audio data, anonymisation may involve altering a voice to mask the speaker's identity while leaving the content of the audio and other speech attributes intact.[234] It may also involve altering attributes of the data that can be linked to, infer, or reveal an individual's identity, such as background noise or the audio content.[235] Audio anonymisation methods include noise addition, voice conversion, or 'disentangled representation' machine learning.[236]
Audio	Any sound that can be heard by the human ear within the acoustic range.[237]
Audio data	Raw or processed electrical signal that is captured and stored in the form of sound, including speech, music, or ambient sound.
Audio data file	A digital vessel for housing audio data. Each audio data file may include the original copy of the audio data, any duplicates, and the audio's metadata.
(Audio) signal	An electrical representation of sound waves that carries the frequency, amplitude, phase, and other essential information of the sound, such that it can be perceived and reproduced. By converting the sound wave into an electrical signal, the audio signal can be captured, transmitted, stored, and processed by electronic devices.[238]
Audit trail	A reproducible step-by-step record of the collection effort and the data's chain of custody.
Authenticity (of evidence)	Whether a piece of evidence is what it professes to be in origin or authorship.

[232] OHCHR, Berkeley Protocol, Glossary.
[233] OHCHR, Berkeley Protocol, Glossary.
[234] Natalia Tomashenko et al, 'The VoicePrivacy 2020 Challenge: Results and findings' (2022) 74 Computer Speech & Language, page 2.
[235] Natalia Tomashenko et al, 'The VoicePrivacy 2020 Challenge: Results and findings' (2022) 74 Computer Speech & Language, page 2.
[236] Natalia Tomashenko et al, 'The VoicePrivacy 2020 Challenge: Results and findings' (2022) 74 Computer Speech & Language, page 2.
[237] Hasan Fayyad-Kazan et al, Verifying the Audio Evidence to Assist Forensic Investigation, (2021) 14(3) Computer and Information Science 25, page 27.
[238] Workplace Technology Term Dictionary: Audio Signal (C&C Technology Group).

Chain of custody	The chronological documentation of the sequence of custodians of a piece of information or evidence, and documentation of the control, date and time, transfer, analysis and disposition of the information.[239]
Collection effort	The entire process of data collection, processing, and data transfer.
Collector	A civil society entity engaged in the collection of data first-hand or from third party sources, where the data could serve as potential evidence of violations of international law.
Critical listening	The aural review of audio data to determine the audio contents and characteristics.
Cryptography	The practice of digitally encoding or decoding information, such as through encryption.[240]
Cryptographic hash value	A fixed-size numerical value generated by an algorithm and attached to a piece of data that provides a mathematical demonstration of the data's uniqueness without revealing the original content.
Cryptographic signature	Two numerical keys (i.e. digital codes) attached to a piece of data, one private and one public, that are generated by an algorithm and mathematically linked. The public key is used to encrypt the data and the private key to decrypt it.
Data set	A collection of data.
Data subject	An identified or identifiable natural person whose personal data is collected. An identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person.[241]
Deletion	The permanent and irreversible removal of the data, including all duplicates and backups.
Due diligence	The exercise of reasonable care and investigation. In a human rights context, it involves 'assessing actual and potential human rights impacts, integrating and acting upon the findings, tracking responses, and communicating how impacts are addressed'.[242]
Duplicate copy	An identical (bit for bit) copy of an original digital asset.[243]
Encryption	The process of making data inaccessible without a decryption key.[244]
Enhancement	The processing steps taken to improve the overall quality and intelligibility of the audio, including reducing unwanted noise or increasing desired sounds, such as speech, without adding to or detracting from the audio's content.[245]
Evidence	Information that has been tendered (formally submitted) to a court as part of a criminal or civil trial for the purpose of proving or disproving an alleged violation of the law.
Evidentiary value	The degree of quality or usefulness for the purpose of proving or disproving an alleged violation of the law. Information that will not be tendered as evidence may still have evidentiary value if it could be used to further an investigation.
Exculpatory	That which may indicate an alleged perpetrator's innocence.
Equality of arms	A legal principle that requires all parties involved in litigation to receive an equal balance of procedural opportunity, as part of the right to a fair trial.

[239] OHCHR, Berkeley Protocol, p. 61.
[240] OHCHR, Berkeley Protocol, Glossary.
[241] GDPR, Art 4(1).
[242] OHCHR, Guiding Principles on Business and Human Rights, Guiding Principle 17.
[243] NIST IR 8387, Digital Evidence Preservation Considerations for Evidence Handlers (2022), p. 6.
[244] OHCHR, Berkeley Protocol, Glossary.
[245] Scientific Working Group on Digital Evidence, SWGDE Best Practices for the Enhancement of Digital Audio (2020), page 3; Hasan Fayyad-Kazan et al, Verifying the Audio Evidence to Assist Forensic Investigation, (2021) 14(3) Computer and Information Science 25, page 29.

Inculpatory	That which may indicate the guilt of an alleged perpetrator.
Informed Consent	The consent provided by a data subject regarding the future use of their data, once the data subject has received clear and understandable information, including: why the data is being sought; how the data will be used and foreseeable consequences; any risk related to the data's future use, and any relevant safeguards in place; the data subject's right to refuse consent and to revoke consent if granted; and, that their consent must be given voluntarily without coercion.[246]
Interoperability	The ability of different information systems, software, or devices to exchange, interpret, and use information seamlessly.
Labelling	The use of a labelling scheme to name a data file, denote the content of the data, and organise a data set to make it easier to search. A label may be a unique yet uniform identification number, or it may comprise the time, date, and/or location of collection. Labelling may also involve critical listening and analysis of the audio data's content, e.g., 'possible troop movements', or '#possiblegunshot'.
Metadata	Data about data; metadata often includes a data file's characteristics and history, and describe how, when, and by whom a digital file was collected, created, accessed, enhanced, modified, and/or formatted.[247] Metadata can be either part of the data upon collection (known as **embedded**, internal, or application metadata), or added to the data (known as **associated**, custom, external, attached, or process metadata) after its collection. Metadata may include: the audio data file label; date and time of creation/of collection; details of the device used to record the audio data (i.e. make, model, serial number, source of power); details of any device(s) used to store the audio data; the technical specifications to which any collection tools are calibrated (a.k.a. 'recorder' metadata); the length of the audio data; the size of the audio data; the geo-location from where the audio data was emitted; the physical location in which the collection was made.
Non-State Armed Group (NSAG)	*Also* 'armed non-State actors': While absent a uniform legal definition, the definition relied upon by civil society group Geneva Call is: '[O]rganized armed entities that are primarily motivated by political goals, operate outside effective state control, and lack legal capacity to become party to relevant international treaties. This includes armed groups, *de facto* governing authorities, national liberation movements, and non- or partially internationally recognized states.'[248]
Original copy	*Also* 'first copy': The unprocessed form of the audio data, as collected. There may be multiple duplicates of the original copy of the audio data, verifiable in that all original copies should share the same content-based hash value or digital signature.
Personal data	Any information relating to an identified or identifiable natural person.[249] A person can be 'identified or identifiable, directly or indirectly, by means reasonably likely to be used related to that data, including where an individual can be identified from linking the data to other data or information reasonably available in any form or medium. If you are using publicly available data, note that this data can also be personal, and therefore may involve some of the same considerations as non-public personal data.'[250]
Personnel	Staff and contractors affiliated with the Collector.

[246] PILPG, Handbook on Civil Society Documentation of Serious Human Rights Violations: Principles and Best Practices (2016), page 9.
[247] OHCHR, Berkeley Protocol, Glossary.
[248] Pascal Bongard & Jonathan Somer, 'Monitoring Armed Non-State Actor Compliance With Humanitarian Norms. A Look at International Mechanisms and the Geneva Call Deed of Commitment', (2011) 93(883) International Review of the Red Cross 673, page 674, fn. 3.
[249] GDPR, Art 4(1).
[250] The UN Global Pulse Risk, Harms and Benefits Assessment Tool, page 3.

Preservation	The policies and strategies required to manage and maintain digital information with enduring value over time, so that the digital information is accessible and usable by its intended users in the future.[251]
Prejudicial effect (of evidence)	The effect caused when the nature of a piece of evidence can have a negative effect on the fairness of the trial as a whole, and in particular, the rights of the accused.
Processing	The phase of the Collection effort that encompasses storage and preservation of the data, any duplication and enhancement of the data, and any analysis of the data.
Probative value	The quality or function of demonstrating or proving the existence of a fact.
Redaction	In the context of audio data, the act of obscuring data with a sound or removing/muting a portion of the audio.
Relevance (of evidence)	An item of evidence is relevant when it pertains to the matters considered at trial, i.e., when it can be used to show that the existence of a particular fact is more or less probable.
Reliability (of evidence)	An item of evidence is considered reliable if the veracity and accuracy of its content can be trusted.
Scraping	A method of extracting mass quantities of data from websites.[252]
Third-party recipients	Organisations or entities with whom the Collector shares audio data.
Third-party sources	Individuals, organisations, or entities that provide audio data to the Collector.
Weight (of evidence)	The relative importance attached to an item of evidence in deciding whether a certain issue has been proven or not.

[251] OHCHR, Berkeley Protocol, Glossary.
[252] OHCHR, Berkeley Protocol, Glossary.