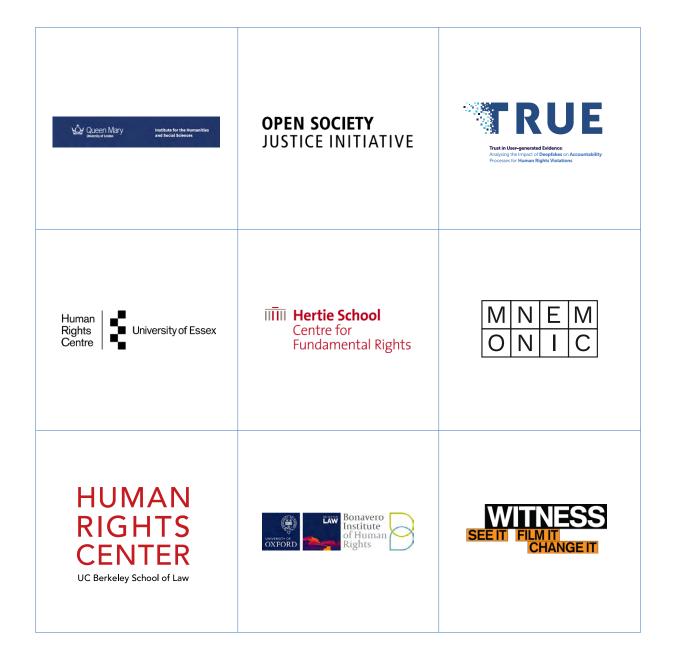


Evaluating digital open source imagery: A guide for judges and fact-finders

## Evaluating digital open source imagery: A guide for judges and fact-finders



Evaluating digital open source imagery: A guide for judges and fact-finders

## **Contents**

About the authors	6
Cite as	6
Introduction	7
What is digital open source information and how to approach its evaluation?	10
What are the distinctive features of digital open source information to be aware of?	11
Key issues to consider when evaluating digital open so	urce
information	12
A. Content information	14
B. Metadata	15
C. Source information	19
D. Location information	20
E. Time information	26
Conclusion	30
Glossary	31
Acknowledgements	33

### About the authors

This guide was prepared following a workshop hosted by the Centre for Fundamental Rights at the Hertie School in Berlin on 29–30 June 2022, and funded by the Digital Verification Unit at the University of Essex. The following individuals (listed here in alphabetical order) designed, drafted, and edited this guide:

Professor Başak Çalı, Director of the Centre for Fundamental Rights and Professor of International Law, Hertie School and Head of Research at the Bonavero Institute of Human Rights and Professor of International Law, University of Oxford.

**Joseph Finnerty,** PhD Candidate, Centre for Fundamental Rights, Hertie School.

**Lindsay Freeman,** Director of Technology, Law, and Policy, Human Rights Center, University of California, Berkeley, School of Law.

**Dr. Alexa Koenig,** Adjunct Professor and Co-Director, Human Rights Center, University of California, Berkeley, School of Law.

Libby McAvoy, Legal Advisor, Mnemonic.

Professor Yvonne McDermott Rees,

Professor of Law, Hillary Rodham Clinton School of Law, Swansea University.

**Dr. Daragh Murray,** Senior Lecturer, School of Law, IHSS Fellow, Queen Mary University of London.

Jana Sadler-Forster, Senior Managing Strategic Litigation Officer, Open Society Justice Initiative and Barrister, Blackstone Chambers.

**Raquel Vazquez Llorente,** Associate Director, Technology, Threats and Opportunities, WITNESS..

**Sarah Zarmsky**, PhD Candidate and Assistant Lecturer, School of Law and Human Rights Centre, University of Essex.

#### Cite as

**Evaluating digital open source imagery:** A guide for judges and fact-finders (2024), published online at <a href="https://www.trueproject.co.uk/osquide">www.trueproject.co.uk/osquide</a>, 2024.

### Introduction

Digital open source information – that is, information that is publicly accessible on the internet¹ – is increasingly used as evidence before domestic and international courts, human rights bodies, and fact-finding bodies,² where it has proven valuable in a variety of contexts.³ For example, open source information has been submitted as evidence in a number of cases before the International Criminal Court,⁴ and videos found online played a significant role in the arrest warrants issued by the Court for Mahmoud Mustafa Busayf Al-Werfalli.⁵ In a first for the European Court of Human Rights, the applicants in *Ponomarenko and Others v. Ukraine and Russia* submitted an interactive digital platform to present open source information.⁶ The case of *Ukraine and the Netherlands v. Russia* also discussed how open source information could be taken into consideration.⁶ Photos and videos from social media have also become instrumental to the findings of United Nations-mandated investigative missions,⁶ and the domestic prosecution of international crimes.⁶

As a relatively new form of evidence, digital open source information may be

- As it is the information most likely to be received by courts in the near future, this document focuses on digital open source imagery, incorporating images and videos, such as satellite imagery, social media posts, or videos taken by a witness on a smartphone. For a full definition of open source information, see Human Rights Center, University of California, Berkeley/UN Office of the High Commissioner for Human Rights, Berkeley Protocol on Digital Open Source Investigations (hereafter, 'Berkeley Protocol') <a href="https://www.ohchr.org/sites/default/files/2022-04/OHCHR\_BerkeleyProtocol.pdf">https://www.ohchr.org/sites/default/files/2022-04/OHCHR\_BerkeleyProtocol.pdf</a>>, 5–8.
- For the purposes of this document, 'human rights bodies' is understood broadly and may include, for example, UN treaty bodies, or Special Procedures of the UN Human Rights Council.
- 3 See for example Sam Dubberley, Alexa Koenig, and Daragh Murray (eds), Digital Witness: Using Open Source Information for Human Rights Investigation, Documentation, and Accountability (OUP 2019); Karolina Aksamitowska, 'Digital Evidence in Domestic Core International Crimes Prosecutions: Lessons Learned from Germany, Sweden, Finland and The Netherlands' (2021) 19 JICJ 189–211; Sarah Zarmsky, 'Why Seeing Should Not Always Be Believing: Considerations Regarding the Use of Digital Reconstruction Technology in International Law' (2021) 19 JICJ 213–225; Alexa Koenig and Ulic Egan, 'Power and Privilege: Investigating Sexual Violence with Digital Open Source Information' (2021) 19 JICJ 55–84.
- This include videos: Prosecutor v Gbagbo and Blé Goudé (Transcript) ICC-02/11-01/15-T-117 (7 February 2017); Prosecutor v Al Mahdi (Judgment and Sentence) ICC-01/12-01/15-171 (27 September 2016); Facebook posts: Prosecutor v Bemba et al. (Decision on 'Prosecution's Fifth Request for the Admission of Evidence from the Bar Table') ICC-01/05-01/13-1524 (14 December 2015); Prosecutor v Bemba et al. (Prosecution's Fifth Request for the Admission of Evidence from the Bar Table) ICC-01/05-01/13-1498 (30 November 2015), §§ 17-18; Prosecutor v Yekatom and Ngaïssona (Transcript) ICC-01/14-01/18-T-023 (29 March 2021), 69; images: Prosecutor v Said (Transcript) ICC-01/04-01/21-T-004 (12 October 2021), 17, and satellite imagery: Prosecutor v Al Hassan (Transcript) ICC-01/12-01/18-T-027 (21 September 2020).
- 5 Prosecutor v Al-Werfalli (Warrant of Arrest) ICC-01/11-01/17-2 (15 August 2017), §§ 11-22; Prosecutor v Al-Werfalli (Second Warrant of Arrest) ICC-01/11-01/17-13 (5 July 2018), §§ 17-18. See further, Emma Irving, 'And So It Begins... Social Media Evidence in an ICC Arrest Warrant' (Opinio Juris, 17 August 2017) <a href="https://opiniojuris.org/2017/08/17/and-so-it-begins-social-media-evidence-in-an-icc-arrest-warrant/">https://opiniojuris.org/2017/08/17/and-so-it-begins-social-media-evidence-in-an-icc-arrest-warrant/</a>>.
- 6 Ponomarenko and Others v. Russia, ECtHR, App. No. 60372/14. Pending. The platform is available at: <a href="https://">https://</a> ilovaisk.forensic-architecture.org/>.
- 7 Ukraine and the Netherlands v. Russia, Admissibility Decision, ECtHR, App. Nos. 8019/16, 43800/14, 28525/20, 30 November 2022, § 472.
- For examples, see Daragh Murray, Yvonne McDermott, and Alexa Koenig, 'Mapping the Use of Open Source Research in UN Human Rights Investigations' (2022) 14 *Journal of Human Rights Practice* 554–581.
- Court of Appeal in The Hague, Case No 22/001283-21 (6 December 2022); Court of Appeals for Western Sweden, Chief Prosecutor v Hassan Mostafa Al-Mandlawi and Al Amin Sultan (Judgment, 30 March 2016); Södertörn District Court, Prosecutor v Mouhannad Droubi (Judgment, 26 February 2015); Örebro District Court, Prosecutor v Saeed (Judgment, 19 February 2019); District Court of The Hague, Case Nos 09/748012-19 and 09/748012-19-P (Judgment, 29 June 2021); District Court of The Hague, Case No 09/74801-19 (Judgment, 16 July 2021).

unfamiliar to many legal professionals. Accordingly, this document provides an overview of key digital open source investigative techniques in order to assist judges and fact-finders in their own evaluation of digital open source information, when it has been submitted by a party to the proceeding or a third party, or obtained via an external report. Importantly, this document does not address how to conduct open source investigations. The sole purpose of this guide is to assist in the evaluation of the credibility, reliability, and probative value of open source information. Certain open source investigative techniques are explained, but only to provide insight into the investigative process.

A framework for how to conduct digital open source investigations is extensively provided for by the Berkeley Protocol on Digital Open Source Investigations ('The Berkeley Protocol').<sup>12</sup> The Berkeley Protocol establishes the professional standards that should be applied in the identification, collection, preservation, analysis and presentation of digital open source information in international criminal and human rights investigations. It includes international standards for conducting online research into alleged violations of international criminal, humanitarian, and human rights law. It also provides guidance on methodologies and procedures for gathering, analyzing and preserving digital information in a professional, legal and ethical manner.

This document builds on the Berkeley Protocol to support judges and other fact-finders in their evaluation of open source information. Because of their importance to accountability and justice mechanisms, this document focuses only on digital open-source imagery (incorporating images and videos), and is consistent with the definitions, principles and techniques described in the Berkeley Protocol.

This guide is organized around a number of key issues that a court or fact-finding body may need to address in their evaluation of open source information, including determining the authenticity of the digital image, and analyzing relevant metadata, source, location, and time information. For each issue, the guide defines relevant terms and techniques, and provides examples in order to inform judges and fact-finders' own evaluative process. In each section, there is a 'Key Takeaways' box, which provides a summary of the information for quick reference. A glossary of relevant technical terms is also included, and terms included in the glossary are hyperlinked and highlighted in bold font.

Different jurisdictions will differ in their rules of admissibility, and in whether expert evidence is required, and, if so, what kind of expertise. This will also be highly

For the purposes of this guide, the terms 'digital open source evidence' and 'open source evidence' may be used interchangeably, solely for readability purposes. The focus is, however, explicitly on digital open source image-based evidence.

For an overview of investigative techniques, see: Dubberley *et al.*, *supra* note 3. For courses or information on open source investigations see: Amnesty International, 'Online Course on Open Source Human Rights Investigations' <a href="https://advocacyassembly.org/en/partners/amnesty">https://advocacyassembly.org/en/partners/amnesty</a>; Institute for International Criminal Investigation, 'Open Source Investigations Course' <a href="https://iici.global/course/open-source-investigation-foundational/">https://iici.global/course/open-source-investigation-foundational/</a>.

<sup>12</sup> The Berkeley Protocol was developed by the Human Rights Centre, University of California Berkeley and the UN Office of the High Commissioner for Human Rights, and involved a series of consultations with international experts.

dependent on the specific facts of the case. This guide is intended to assist with the assessment of any submitted material, so that accountability mechanisms can capitalize on digital open source information's full potential. It is our belief that open source information will continue to be invaluable to the pursuit of accountability. The guide is intended to be illustrative and not exhaustive. Prominent open source investigative techniques are addressed, but new techniques emerge continually.

**Key Takeaway:** This guide is intended to assist judges and other decision makers in their assessment of open source information, by explaining some of the most common open source investigative techniques. It is not a guide on how to conduct open source investigations.

# What is digital open source information and how to approach its evaluation?

The Berkeley Protocol defines <u>open source information</u> as 'information that any member of the public can observe, purchase or request, without requiring special legal status or unauthorized access'.<sup>13</sup> <u>Digital open source information</u> is 'publicly available information in digital format, which is generally acquired from the internet'.<sup>14</sup> In an accountability context, digital open source information consists most notably of social media posts, images, videos, documents and audio recordings on the internet, satellite imagery, and government-published data.

For judges or other fact-finders, there are a number of factors to consider when evaluating the evidentiary value of digital open source information. 'Verification' refers to the assessment of all available information associated with the material. The process of verifying open source information involves a combination of different techniques, such as geolocation, chronolocation, and metadata analysis, and is not limited to one specific technique. When evaluating open source information it is important to examine the investigative methodology employed.

**Key Takeaways:** The evaluation of open source information is centered around ensuring that an appropriate verification process has been conducted. The actual verification techniques used will inevitably differ on a case-by-case basis. It may be useful to bear in mind that each technique is part of a corroborative puzzle, and investigators should rule out alternative possibilities.

<sup>13</sup> Berkeley Protocol, § 1.

<sup>14</sup> lc

These techniques are discussed further below in Section 4.

# What are the distinctive features of digital open source information to be aware of?

For the most part, digital open source imagery should be approached in the same way as any other form of evidence, considering existing factors such as corroboration and source reliability. However, there are a few additional key considerations to be borne in mind:

- First, open source imagery may not include traditional indicia of authenticity, such as information on the person who recorded the material, or details about the original device on which the footage was recorded. Importantly, a user may post content that they themselves did not record.
- Second, as is often the case with social media accounts, the individual(s)
  associated with an account may be anonymous or unknown. For example,
  some social media platforms do not require users to provide their real
  name, may allow users to change their username repeatedly, and/or
  multiple individuals may post to a single account.
- Third, the nature of digital environments allows for high volumes of material to spread quickly, and the person who posted the material for the first time may be unknown.
- Fourth, as discussed in Section 4, content may be inauthentic in a variety of ways. Tools for generating or editing content are now much more accessible, and can be used without professional training or complex software.
   Unlike physical evidence, digital content may be tampered with remotely.

**Key Takeaway:** In general, digital open source information should be approached in the same manner as any other form of evidence. However, there are few unique attributes that are worth paying attention to. Proponents of the evidence should address questions that arise from the ways in which open source information is characteristically different from other forms of evidence.

## Key issues to consider when evaluating digital open source information

This section identifies the key issues to consider when evaluating the authenticity and reliability of <u>digital open source information</u>. In an open source investigations context, <u>verification</u> is the process by which the accuracy and validity of information is assessed. Digital imagery is deemed to be authentic and reliable once it has been demonstrated that it represents what it is claimed to represent. When evaluating how digital imagery has been verified and authenticated, consider whether and how the investigator's analysis assesses: (A) the content of the imagery itself, (B) the metadata, (C) the source, (D) the location, and (E) the time.

There are a wide range of reasons why online content may not be what it is purported to be.<sup>16</sup> These include, but are not limited to, the following:

- Misattribution of place, time or decontextualization: Even though content may depict real events, it is possible that the time or place of a photo or video is misattributed or that the content is taken out of context. For example, a video allegedly depicting Turkish attacks in northern Syria was circulated across multiple major news outlets in 2019. However, shortly after, it was found that the video was actually from a gun range in Kentucky in the United States of America.<sup>17</sup>
- Edited content (shallowfakes): In some instances, edited photos or videos may be presented as original content. They may be cut, have filters applied, elements may be added or deleted, or the video frames sped up or slowed down (such edited content is otherwise known as 'shallowfakes'). An example of this is a real video of Nancy Pelosi, the Speaker of the House in the USA, that was edited to make it appear that she was intoxicated and slurring her words. The video was later debunked.

<sup>16</sup> Claire Wardle, 'Fake news. It's complicated.' (*First Draft News*, 16 February 2017) < <a href="https://firstdraftnews.org/articles/fake-news-complicated/">https://firstdraftnews.org/articles/fake-news-complicated/</a>>.

Heather Murphy, 'ABC Apologizes for Showing Video from U.S. Gun Range in Report on Syria' (*The New York Times*, 14 October 2019) <a href="https://www.nytimes.com/2019/10/14/business/media/turkey-syria-kentucky-gun-range.html">https://www.nytimes.com/2019/10/14/business/media/turkey-syria-kentucky-gun-range.html</a>>.

<sup>18</sup> Ashley Stoll, 'Shallowfakes and Their Potential for Fake News' (Washington Journal of Law, Technology & Arts, 13 January 2020) <a href="https://wjlta.com/2020/01/13/shallowfakes-and-their-potential-for-fake-news/">https://wjlta.com/2020/01/13/shallowfakes-and-their-potential-for-fake-news/</a>>.

Hannah Denham, 'Another fake video of Nancy Pelosi goes viral on Facebook' (Washington Post, 3 August 2020) <a href="https://www.washingtonpost.com/technology/2020/08/03/nancy-pelosi-fake-video-facebook/">https://www.washingtonpost.com/technology/2020/08/03/nancy-pelosi-fake-video-facebook/</a>>.

#### Modified metadata:

- » Automatically modified or deleted <u>metadata</u>: The metadata attached to content may be automatically modified by a platform when that content is uploaded. For example, WhatsApp – like most social media and digital communication platforms – removes most metadata from content uploaded to the platform.
- » Manually modified or deleted metadata: The metadata attached to content may be modified, knowingly or not, and may show an incorrect location, recording device, or timestamp. It may also be fully or partially erased. Modification can be done through a metadata editor or through a built-in function on certain operating systems, and it may be undertaken for different purposes.<sup>20</sup>
- Staged content: Content may be staged using actors and film or television sets. An example of this occurred in 2014 in relation to the conflict in Syria, when a video of a young boy rescuing a girl under gunfire titled 'Syrian Hero Boy' and initially presented as authentic went viral.<sup>21</sup> It was later revealed that a group of filmmakers were behind the video, which was actually not from the Syrian conflict but was filmed with actors on a set in Malta.
- Al-generated or manipulated content (deepfakes or synthetic media):

  As artificial intelligence (AI) technology becomes more widely accessible, digital open-source audio, photos, and videos may be generated or edited by AI.<sup>22</sup> Deepfakes are an example of AI-enabled techniques for synthetic media generation. They are a new form of audiovisual manipulation that allows people to create realistic simulations of someone's face, voice or actions. For example, a deepfake video of Ukrainian President Zelensky calling his troops to surrender was circulated in 2022 on social media.<sup>23</sup> Synthetic media technology also enables users to add or remove objects easily, alter background conditions, create an image of a person who does not exist, or generate an image of an event or object from a text description, among other features.<sup>24</sup> AI-generated or edited content can be difficult to detect and may require analysis by an expert on AI synthesis or media forensics. Tools that claim to identify deepfakes are not always

Modification may be undertaken for misleading purposes, or for other reasons; in some instances, for example, redaction of the metadata may be needed to preserve anonymity. For information on metadata editing processes, see Casey Schmidt, 'Revamp your information with these unique metadata editors' (Canto, 2 February 2021) <a href="https://www.canto.com/blog/metadata-editor/">https://www.canto.com/blog/metadata-editor/</a>; Mauro Huculak, 'How to edit image metadata on Windows 10' (Windows Central, 10 January 2017) <a href="https://www.windowscentral.com/how-edit-picture-metadata-windows-10">https://www.windowscentral.com/how-edit-picture-metadata-windows-10</a>>.

<sup>21</sup> BBC News, '#BBCTrending: Syrian 'hero boy' video faked by Norwegian director' (*BBC News*, 14 November 2014) <a href="https://www.bbc.com/news/blogs-trending-30057401">https://www.bbc.com/news/blogs-trending-30057401</a>>.

<sup>22</sup> WITNESS, *Deepfakes* (2022), available online at <a href="https://www.mediafire.com/file/421ov54c77t04tq/Backgrounder\_Deepfakes\_2022.pdf/file>">https://www.mediafire.com/file/421ov54c77t04tq/Backgrounder\_Deepfakes\_2022.pdf/file>">https://www.mediafire.com/file/421ov54c77t04tq/Backgrounder\_Deepfakes\_2022.pdf/file>">https://www.mediafire.com/file/421ov54c77t04tq/Backgrounder\_Deepfakes\_2022.pdf/file>">https://www.mediafire.com/file/421ov54c77t04tq/Backgrounder\_Deepfakes\_2022.pdf/file>">https://www.mediafire.com/file/421ov54c77t04tq/Backgrounder\_Deepfakes\_2022.pdf/file>">https://www.mediafire.com/file/421ov54c77t04tq/Backgrounder\_Deepfakes\_2022.pdf/file>">https://www.mediafire.com/file/421ov54c77t04tq/Backgrounder\_Deepfakes\_2022.pdf/file>">https://www.mediafire.com/file/421ov54c77t04tq/Backgrounder\_Deepfakes\_2022.pdf/file>">https://www.mediafire.com/file/421ov54c77t04tq/Backgrounder\_Deepfakes\_2022.pdf/file>">https://www.mediafire.com/file/421ov54c77t04tq/Backgrounder\_Deepfakes\_2022.pdf/file>">https://www.mediafire.com/file/421ov54c77t04tq/Backgrounder\_Deepfakes\_2022.pdf/file>">https://www.mediafire.com/file/421ov54c77t04tq/Backgrounder\_Deepfakes\_2022.pdf/file>">https://www.mediafire.com/file/421ov54c77t04tq/Backgrounder\_Deepfakes\_2022.pdf/file>">https://www.mediafire.com/file/421ov54c77t04tq/Backgrounder\_Deepfakes\_2022.pdf/file>">https://www.mediafire.com/file/421ov54c77t04tq/Backgrounder\_Deepfakes\_2022.pdf/file>">https://www.mediafire.com/file/421ov54c77t04tq/Backgrounder\_Deepfakes\_2022.pdf/file>">https://www.mediafire.com/file/421ov54c77t04tq/Backgrounder\_Deepfakes\_2022.pdf/file>">https://www.mediafire.com/file/421ov54c77t04tq/Backgrounder\_Deepfakes\_2022.pdf/file>">https://www.mediafire.com/file/421ov54c77t04tq/Backgrounder\_Deepfakes\_2022.pdf/file>">https://www.mediafire.com/file>">https://www.mediafire.com/file>">https://www.mediafire.com/file>">https://www.mediafire.com/file>">https://www.mediafire.com/file>">https://www.mediafire.com/file>">https://www.mediafire.com/file>">https://www.mediafire.com/file>">h

<sup>23</sup> Bobby Allyn, 'Deepfake video of Zelenskyy could be 'tip of the iceberg' in info war, experts warn' (NPR, 16 March 2022) <a href="https://www.npr.org/2022/03/16/1087062648/deepfake-video-zelenskyy-experts-war-manipulation-ukraine-russia?t=1660657155956">https://www.npr.org/2022/03/16/1087062648/deepfake-video-zelenskyy-experts-war-manipulation-ukraine-russia?t=1660657155956></a>.

<sup>24</sup> Open AI, 'DALL:E: Creating Images from Text' (OpenAI, 5 January 2021) <a href="https://openai.com/blog/dall-e/">https://openai.com/blog/dall-e/</a>.

accurate and should not be solely relied upon when examining suspicious content; context and corroboration should also be considered.<sup>25</sup> Determining when a piece of content was created can give insight into any media generation or modification tools that were available at the time.

Nevertheless, even edited or inauthentic content may have evidentiary value.<sup>26</sup> Content may be manipulated without the intention to mislead. For example, a video may be cut and joined to another video, without the editor intending to suggest that the two parts ran on from each other sequentially. Or, even if intended to mislead, there may still be aspects of the photo or video that have probative value, such as the date or time that the footage was recorded, or the content itself if it is propaganda. This may also speak to other factors such as the mental elements of a crime (mens rea). This type of information should of course be approached with appropriate caution. Normal considerations for analyzing digital open source information as outlined in the Berkeley Protocol should be applied by investigators in order to attribute the appropriate value, if any, to potentially inauthentic digital imagery.

**Key Takeaways:** Online open source imagery may not always be what it is purported to be, for multiple reasons, including: misattribution, editing, modification of metadata, staging, and the use of artificial intelligence to create or manipulate content. Many of these can be identified using appropriate verification techniques. Inauthentic digital imagery may still have evidential value.

#### A. Content information

Although open source information may be presented in a different manner to more traditional forms of photo or video evidence, the content (i.e. the information depicted in the photo or video) should be analyzed in the same way. When evaluating an investigator's report, two components should be considered.

**First, the process followed when the content was examined.** The investigator should follow professional standards for the collection, analysis, and preservation of open source evidence as outlined in the <u>Berkeley Protocol</u>. They should also be transparent as regards any known biases or limitations of their work, and should have attempted to offset both cognitive and technical biases where possible.<sup>27</sup> The investigator should state whether they tested alternative hypotheses or considered other methods for interpreting or challenging their work.

<sup>25</sup> Sam Gregory, 'The World Needs Deepfake Experts to Stem This Chaos' (Wired, 24 June 2021) < <a href="https://www.wired.com/story/opinion-the-world-needs-deepfake-experts-to-stem-this-chaos/">https://www.wired.com/story/opinion-the-world-needs-deepfake-experts-to-stem-this-chaos/</a>>.

See, e.g., Prosecutor v. Nahimana, Barayagwiza & Ngeze, Judgment, International Criminal Tribunal for Rwanda, Case No. ICTR-99-52-T, 3 December 2003, § 274.

<sup>27</sup> Yvonne McDermott, Alexa Koenig, and Daragh Murray, 'Open Source Information's Blind Spot: Human and Machine Bias in International Criminal Investigations' (2021) 19 JICJ 85–105.

Second, whether the investigator's analysis and findings are appropriate to their expertise. For instance, in some cases it may be necessary for an investigator to consult a technical, subject matter, or industry expert (such as forensic pathologists, botanists, or a medical, weapons, military, or geospatial expert). In other cases it may be necessary for an investigator to consult individuals with context-specific knowledge, as investigators without the appropriate expertise in the context or subject matter depicted – such as those who are personally unfamiliar with an area allegedly depicted – may miss context clues that plainly disprove findings or fail to appropriately interrogate biases, assumptions or mis- and dis-information. The presenter of the evidence should consult experts when necessary and not make claims about what is depicted in the imagery that fall outside of their expertise. If the language of the original source differs from the language of the report, accuracy of translation should be considered.

**Key Takeaways:** Open source content is analyzed in the same way as traditional photos or videos, with two points for particular scrutiny. First, an investigator's process should be assessed to ensure that they exercised due diligence when analyzing the content. Second, an investigator's findings should be appropriate to their knowledge and expertise.

#### B. Metadata

<u>Metadata</u> is data that describes and gives information about specific pieces of content, such as the photo or video that is being assessed.<sup>28</sup> There are two principal sets of possible metadata for each item: metadata attached at the time of creation, editing, or distribution; and metadata added by investigators as part of the analysis or preservation process. Each can provide different information.<sup>29</sup>

## Metadata attached at the time of content creation, editing or distribution

Metadata embedded at the time digital content is created can include the time, date, and location of capture, as well as information such as the type of device on which the content was created. The creation of metadata varies across the type of device that created the content, and depends largely on how the device is configured, or whether the platform it was uploaded to automatically 'strips' (i.e. removes) metadata.

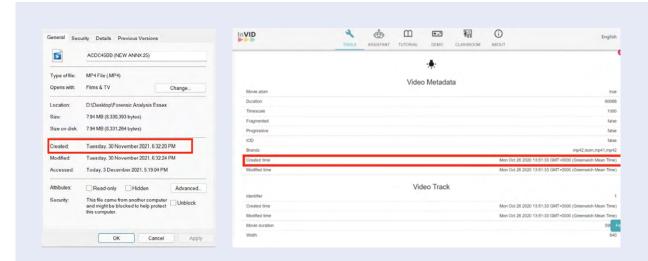
<sup>28</sup> Berkeley Protocol, § 184.

<sup>29</sup> It should be noted that metadata may also be modified or created in other ways. For example, metadata may be deliberately modified post creation, in order to tamper with the time of recording, etc. Equally, metadata may be automatically added if the content is edited, using a photo or video editing package.

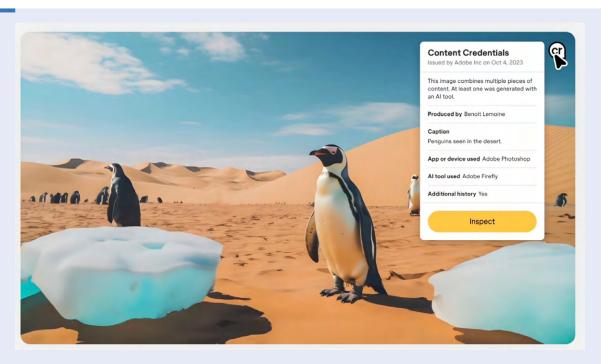
A number of factors can lead to variations in the metadata, such as:

- The timestamp: This may be impacted by whether the device is on a 'default' time zone setting;
- B. The approximate GPS coordinates: These may be affected by factors such as the number and location of cell towers in the vicinity, or the network provider's level of coverage in the area; and
- C. Derivative metadata: Some mobile phones derive altitude on the basis of other metadata points. Any variations in the source metadata will have knock on effects..

In addition, metadata typically needs to be examined using a metadata viewer to extract and interpret it. Depending on which metadata viewer is used, the results may be slightly different, as the following examples illustrate (Figure A). For content generated or edited by AI, some tools may embed details about the software that created or modified the image or audio, or the generative model that was used (Figure B).



**Figure A:** Screenshots of two metadata viewer outputs taken from the Digital Verification Unit at the University of Essex's verification of a video depicting an event at the Lekki Tollgate in Nigeria. The top image shows the metadata extracted using Microsoft File Explorer, listing a creation date and time of 30 November 2021 at 6:32pm. The bottom image shows the metadata extracted using the InVid Toolkit, listing a creation date and time of 26 October 2020 at 13:51. Researchers attributed this discrepancy to the fact that the Microsoft File Explorer metadata (dated 30 November 2021) represented the time in which the file was uploaded to the computer, while the InVid metadata (dated 26 October 2020) was from the actual time of recording.



**Figure B:** Example of content credentials explaining the methods used to create an image using Artificial Intelligence. Source: <a href="https://blog.adobe.com/en/publish/2023/10/10/new-content-credentials-icon-transparency">https://blog.adobe.com/en/publish/2023/10/10/new-content-credentials-icon-transparency</a>

Metadata accuracy can also depend on user setup for devices such as surveillance cameras, in which the time needs to be manually entered and thus can easily be wrong.

Another challenge for open source researchers is missing metadata, since metadata is often 'stripped' from content when it is posted to social media websites or sent through messaging apps such as WhatsApp. As much of the content of relevance to this guide is obtained from social media it is likely that the original metadata will not be attached to any reports submitted by researchers or other actors.

#### Metadata attached at upload or while online

While original file metadata is often removed from online content during the upload process, useful metadata can be added to content at the time of upload and during the time that it exists online. As outlined in subsections D and E below, time and location details recorded with content during the upload process can provide investigators with additional clues for their assessment of this key information. Further, online interactions with the content – such as comments, shares, and more – can provide investigators with useful insight.

In the last few years, numerous tools that embed metadata into an image have emerged. These media provenance standards can track how a piece of media is made, as well as the modifications it may have undergone, in a manner that makes it very hard to tamper with the cryptographic signature without leaving evidence of the attempt. For verification purposes, the most valuable tools are those that

embed metadata that is cryptographically hashed at the point of collection, instead of at a later stage (as the image could have then been manipulated in the interim). These tools are often referred to as 'controlled-capture technology'. The design of this software may vary and the integrity of their metadata should not be taken for granted. Similarly, the fact that an image is lacking cryptographically hashed metadata does not mean the content is unreliable or cannot be authenticated by other means.

#### Investigator-added metadata

Metadata may also be added by investigators, after they obtain the content, as part of the analysis or preservation process. For example, investigators might add some information to the digital package that represents their own interpretations of the content, such as new data about the type of event (e.g. 'air strike' or 'torture'). As part of the preservation process, investigators may also add metadata such as a timestamp (indicating when the investigator received the data or an estimate of the time of the depicted event) or a hash value. A <u>hash value</u> is a unique form of digital identification (an alphanumeric string) that confirms, through the use of cryptography, that the content collected has not been modified since the time the hash was calculated.<sup>30</sup> Hash values may be assigned to an item to help establish that it has not been tampered with from the time the hash was applied to the point it is submitted to a court or other fact-finding body. If a digital image is modified even slightly this will result in an entirely new hash value.

#### Invisible watermarks

For synthetic media, invisible watermarks are embedded at the pixel level of visual content or encoded in audio frequency. They are imperceptible to the human eye or ear, but they can be detected by software trained to spot them. They require technical know-how to be edited or removed. A forensic image analyst may be able to confirm whether a video or image has been created using artificial intelligence.

**Key Takeaways:** Metadata should be viewed as part of an overall corroborative picture. Metadata can be useful for building a hypothesis as to the time or location of a photo or video, or whether content has been edited, but should always be evaluated. Incorrect or missing metadata does not necessarily mean that the content is unreliable. For synthetic media, invisible watermarks can help adequately trained detection tools provide more context about an image.

#### C. Source information

Traditionally, witnesses testify as to the source of a photo or video. However, given the distinct nature of the digital environment, this may not be possible for online open source digital imagery. For instance, the scale of digital content could make it impossible to bring in witnesses for each photo or video, or the uploader or sharer of the content may be anonymous, or deceased. Importantly, most investigative techniques described in this section will not confirm the identity of a photo or video's source. Instead, they are useful in determining key characteristics about the source (such as potential political affiliations, apparent physical location, or a regular connection to an event, etc.) that can facilitate an assessment of the source's reliability.

There are multiple actors whose roles should be considered when evaluating the source of a digital image, including the creator (who recorded the original content), the uploader (who posted the content to the internet), the sharer (who distributed the content either online or through messaging group chats), and the compiler or editor (who may have assembled multiple videos into one or modified the content in some way). These roles may overlap. For instance, the creator may also be the uploader.

In some cases, indicators as to the identity of the uploader may exist. For example, certain social media accounts may be 'verified',<sup>31</sup> suggesting the likely identity of the uploader, although the level of verification associated with different social media platforms may vary significantly. 'Verification' here refers to the poster, not the content. It should not be assumed that content posted by a 'verified' account is *de facto* credible or reliable.

Given the nature of the digital environment, the source may be anonymous or pseudo-anonymous. A pseudo-anonymous account could be an account representing a network of people filming and sharing content, usually pertaining to a certain cause or conflict. An example is 'Raqqa is Being Slaughtered Silently', a group of activists who regularly post about ISIS activities in Raqqa, Syria.<sup>32</sup> In some instances, anonymous accounts can be imitator accounts (those which impersonate a famous individual or entity) or bots (accounts that automatically generate posts). In order to evaluate an anonymous source, it is useful to analyze the behavior of the account. For instance, whether the account regularly posts about a conflict or cause, and whether that content is consistent within the overall context (e.g., contains information about known regions or parties to a conflict), or whether the source demonstrates political affiliation or bias. Another indicator of reliability might be if other verified, credible accounts follow that account.

When an account is verified on a social media platform (such as Facebook, Twitter, Instagram, or YouTube), it means that the platform has confirmed (by their own standards) that the profile is authentic to the person or business it represents. Platforms consider a variety of factors when determining whether to verify an account, including identification using ID or an official email address, news coverage, follower counts and the activity of the account.

<sup>32</sup> See for example the Facebook page 'Raqqa is Being Slaughtered Silently تمرب حبذت قور ل , available at <a href="https://m.facebook.com/Raqqa.SI/?\_tn\_\_=%2Cg">https://m.facebook.com/Raqqa.SI/?\_tn\_\_=%2Cg</a>.

It is, of course, possible that sources that do not regularly post about a situation can nonetheless post a single piece of credible, relevant content. This was the case with respect to a video depicting the killing of two women and two children in Cameroon in 2018, which was shared widely across the media.<sup>33</sup> The verification of this video led to the arrest and ultimate conviction of soldiers involved in the executions.<sup>34</sup>

**Key Takeaways:** There are multiple ways in which the source of a digital image may be evaluated. However, given the nature of the online environment in which footage can be shared and re-shared, and that accounts may be anonymous, it may be difficult to determine the source of an open source digital image or video with absolute certainty. In cases where the source cannot be identified, this does not mean that the content is unreliable or has no probative value.

#### D. Location information

Investigators use a variety of techniques to determine where a photo or video was taken. Geolocation refers to 'the identification or estimation of the location of an object, an activity or the location from which an item was generated'.<sup>35</sup> The geolocation process is intended to determine where the content was created. A geolocation report should include a transparent presentation of the granular pieces of information that were used to identify the geographical location. The precision obtained is dependent upon a number of variables.

The metadata for a photo or video may include a geotag (including GPS coordinates), which can be used as a starting point for assessing where the footage was taken. However, the original metadata may be missing (especially if a photo or video was posted on social media, which removes metadata) or could have been altered, and therefore should only be used as a piece of the overall corroborative picture.

At its simplest, geolocation entails matching geographic characteristics visible within the content of an image (either natural or human-built structures) to an actual place, using satellite imagery or other known reference material, such as Google Street View. Generally, the more unique characteristics that are present and can be matched to reference data, the higher the degree of confidence that the photo or video was taken at that particular location.

<sup>33</sup> Nick Turse, 'Cameroon is a close U.S. ally-and its soldiers carried out a shocking execution of women and children' (*The Intercept*, 26 July 2018) <a href="https://theintercept.com/2018/07/26/cameroon-executions-us-ally/">https://theintercept.com/2018/07/26/cameroon-executions-us-ally/</a>>.

<sup>34</sup> BBC News, 'Cameroon soldiers jailed for killing women and children' (BBC News, 21 September 2020) <a href="https://www.bbc.co.uk/news/world-africa-54238170">https://www.bbc.co.uk/news/world-africa-54238170></a>.

<sup>35</sup> Berkeley Protocol, § 190.

#### Reverse image or video search

A <u>reverse search</u> involves uploading an image or stills from a video to a search engine, so that the search algorithm can identify other copies of the same or similar images on the internet. A reverse image search may reveal if an image or video was posted online prior to the date of its alleged creation. This can demonstrate that an image or video is not what its poster claims it to be. However, a lack of matches does not conclusively demonstrate that the image is credible. For example, earlier images may have been taken offline, or may never have been uploaded before.

The limitation of a reverse image search is that it only scans within a search engine's database, which includes a small percentage of the content on the internet. It does not, for example, include materials on the <u>deep web</u> (which is not indexed to search engines) or the <u>dark web</u> (the part of the Internet that can only be accessed through specialized software, such as the Tor browser). It is possible that a reverse image search at the time of the investigation does not return any results, but the same search process run at a later date – such as at the time of court proceedings – may yield results. Search engine databases are constantly growing to encompass more content, and what is indexed can vary significantly between search engines. As the volume of synthetic media online increases, this may affect audiovisual archives, skewing results. Conversely, Al generated or edited media may still return a reverse image search result.

In some cases, a reverse image search of a photo or a keyframe from a video can yield a direct match to a particular location. This occurs mainly when the image includes a known street, landmark, or other structure that is easily identifiable.

Ultimately, a reverse image search of an item discovered on the internet can be used to assess whether the photo or video being analyzed by the investigator was the first known online posting of the imagery. Reverse image searches can also be used to rule out certain locations if the image has appeared online in a search database previously and has already been confirmed to a location different from that under investigation.



**Figure C:** Photo of the Taj Mahal and screenshot of reverse image search using Google. The photo was uploaded to search using Google's 'search by image' feature and yielded the above result, which provided the name of the landmark.

#### Using clues from the image

To determine where a digital image was taken, there are sometimes clues within the imagery that may be used as lead information. Visible, location-related details such as businesses and street names are searchable. Other commonly visible information may also indicate location: car license plates are typically specific to a location, as are police or military uniforms, while street lamps may also differ significantly from location to location. These types of features in the photo or video can help to narrow down reasonable possibilities for where the content was created.

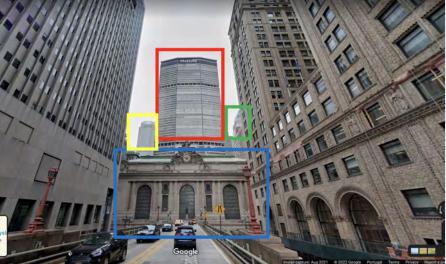
An investigator will often annotate a screenshot of the photo or video with colored boxes in order to highlight certain features, such as distinct buildings, trees, or visible

mountain ranges, which are used as part of the analysis. After these features are identified, investigators may then match them to a location using one or more of the following techniques.

#### Street view or 3D maps

Street view or other 3D maps can be used to match features where available and can be useful when known buildings or landmarks are already labeled.

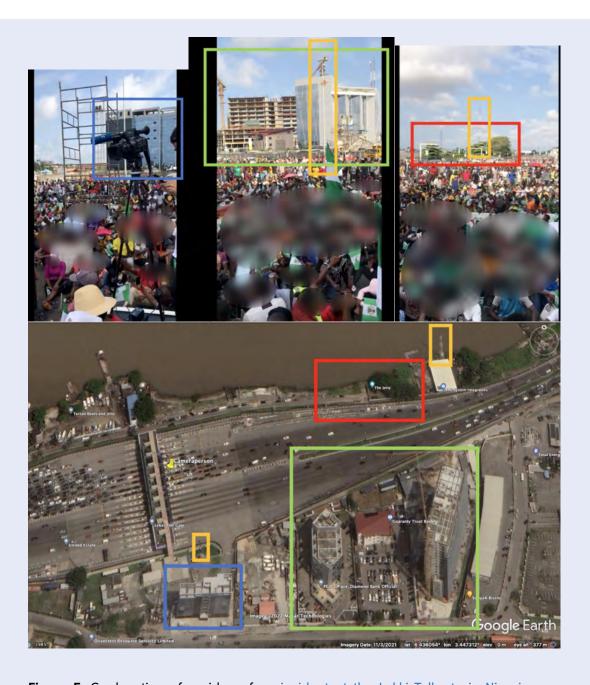




**Figure D:** An example of using the Google Street view feature to match a Google Images photograph of Grand Central Station in New York City (top image) to street view imagery on Google Maps (bottom image).

#### Satellite imagery

Satellite imagery can be used to match features from a bird's eye view. This is typically done using markers, such as colored boxes, to show which structures from a photo or video match to which points on satellite imagery.



**Figure E:** Geolocation of a video of an <u>incident at the Lekki Tollgate in Nigeria</u> undertaken by the Digital Verification Unit at the University of Essex. Investigators annotated the images with colored boxes to indicate where structures seen in the open source video appear to match structures visible in satellite imagery from Google Earth Pro.

#### Terrain mapping

Terrain mapping involves looking for topographic features like mountain ranges and matching them to satellite imagery, which can be useful if no there is no street view or few human-made structures in the captured footage, or only poor quality imagery is available.



**Figure F:** Terrain mapping example taken from Amnesty International's Citizen Evidence Lab's geolocation of a video of an <u>incident in Mahbere Dego</u> in Ethiopia. The mountains identified with red lines in the background of the video were matched to satellite imagery from Google Earth Pro.

#### Challenges of geolocation

<u>Geolocation</u> is often time-consuming and can be very difficult. It is therefore important to understand its limitations. The investigative report accompanying the geolocation should clearly articulate the methodology adopted and explain any limitations of the analysis.

For example, geolocation typically involves comparing the content to satellite imagery, which can vary in availability and quality. Depending on the point on Earth, satellite imagery can be of lower quality due to restrictions (e.g., when satellite imagery companies are prevented by governments from making high resolution imagery available for certain areas, which previously occurred in Gaza).<sup>36</sup> Satellite imagery may also be obstructed by cloud cover, or contain deliberate black outs of areas by governments (e.g., at the time of writing, China has blocked out areas of the Xinjiang region on Baidu Maps).<sup>37</sup> If the footage was filmed indoors, geolocation by comparison to satellite imagery may be impossible; other methods such as reverse image and leads-based searching must be used instead. Due to one or more of these factors, or simply the absence of any defining characteristics in the footage (e.g., a video filmed at sea with no visible landmarks), geolocation may be

<sup>36</sup> Christopher Giles and Jack Goodman, 'Israel-Gaza: Why is the region blurry on Google Maps?' (BBC News, 17 May 2021) <a href="https://www.bbc.co.uk/news/57102499">https://www.bbc.co.uk/news/57102499</a>>.

<sup>37</sup> Alison Killing, Megha Rajagopalan, and Christo Buschek, 'Blanked-Out Spots On China's Maps Helped Us Uncover Xinjiang's Camps' (Buzzfeed News, 27 August 2020) <a href="https://www.buzzfeednews.com/article/alison\_killing/satellite-images-investigation-xinjiang-detention-camps">https://www.buzzfeednews.com/article/alison\_killing/satellite-images-investigation-xinjiang-detention-camps</a>.

very difficult or impossible. In cases where geolocation is not possible, it is important that the investigator provide an explanation for why they were unable to geolocate the content.

**Key Takeaways:** There are multiple methods that investigators can use when determining where a photo or video was taken, including but not limited to: metadata analysis, reverse image and video searches, using leads from the image, matching features on satellite imagery, and terrain-mapping. All have limitations. Reliable location assessments will consider and use multiple methods for analysis.

#### E. Time information

It is generally more difficult to determine when a photo or video was taken than where it was taken, although the investigative process involves similar methods. When evaluating when a photo or video was generated, an investigator will typically first look at the metadata to determine if it contains a creation date timestamp. However, in many cases, the metadata will either be inaccurate or missing, and investigators will need to use other techniques to chronolocate a photo or video. Chronolocation refers to 'the corroboration of the dates and times of the events depicted in a piece of information, usually visual imagery'.<sup>38</sup> The following are some popular chronolocation techniques.

#### Upload time and date stamp

Time and date stamps for online posts, which will always be listed along with a post on social media websites, can also be useful for chronolocation. Importantly, time and date stamps indicate when that specific account uploaded the content, and not when the content was created. As such, time and date stamps can be used as indicators for a timeframe for when an event took place, and can be a useful starting point for chronolocation, but cannot be relied upon to indicate the exact time something happened.

Additionally, the timestamp may differ depending on the platform the content was posted on, as different social media websites operate in different time zones or follow different timestamping protocols. For instance, some platforms may show posts in the viewer's local time zone, or in the time zone where the company is located, and not the time zone for the location where a photo or video was originally taken.

#### Identification of leads within an image

There are sometimes clues within the imagery that may help to determine the date or time a photo or video was taken. This has always been the case but accessing these clues may now be easier. For example, if there are temporary details such as

<sup>38</sup> Berkeley Protocol, § 191.

posters in the background of an image, easily available historical imagery on street view may assist in determining when those posters were put up. This occurred when former Trump aide, George Papadopoulos, was accused of having left the country while under investigation by the FBI after a photo was posted of him in London.<sup>39</sup> Journalists were able to dispute that the image was taken at the alleged time due to clues in the background of the image, including a poster on a lamp post; ultimately it was found that the photo was four years old and not contemporary.<sup>40</sup> Other features in the photo or video, such as the presence of Christmas decorations, could also help to narrow down the season in which the content was created. Of course, such details may be edited into, or out of, the content.

#### Shadow analysis

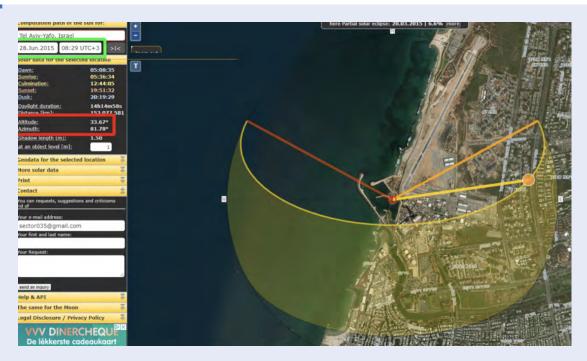
Shadow analysis recognizes that the shadows cast by objects and individuals can indicate the position of the sun at the time of capture. If the location and date are known, the length and direction of any sun-made shadows depicted in the footage can indicate the approximate time at which a photo or video was taken. Shadow calculators independently estimate the approximate shadow length and direction for different sized structures based on the position of the sun at the estimated location, date, and time input by the investigator. However, time analysis using these tools is inexact. Where possible to conduct, shadow analysis provides the investigator's best estimate by their visual examination for a window of time in which a photo or video was taken, but not an exact calculation.







- 39 George Bowden and Jack Sommers, 'London Photo Of George Papadopoulos Was Taken At Least Four Years Ago' (HuffPost, 31 October 2017) <a href="https://www.huffingtonpost.co.uk/entry/george-papadopoulos-twitter-donald-trump-london-picture\_uk\_59f8793ae4b09b5c2568ff4c">https://www.huffingtonpost.co.uk/entry/george-papadopoulos-twitter-donald-trump-london-picture\_uk\_59f8793ae4b09b5c2568ff4c</a>.
- 40 Ibic
- 41 Youri van der Weide, 'Using the Sun and the Shadows for Geolocation' (*Bellingcat*, 3 December 2020) <a href="https://www.bellingcat.com/resources/2020/12/03/using-the-sun-and-the-shadows-for-geolocation/">https://www.bellingcat.com/resources/2020/12/03/using-the-sun-and-the-shadows-for-geolocation/</a>>.



**Figure G:** Chronolocation by Sector035 of a photo taken on the Israel National Trail in Tel Aviv. In the first photo, the shadow was identified, then its length was extended in line with the buildings in the background (second and third photos). SunCalc was then used to estimate the corresponding time. (<a href="https://medium.com/quiztime/lining-up-shadows-2351ae106cec">https://medium.com/quiztime/lining-up-shadows-2351ae106cec</a>)

#### Historical weather analysis

Historical weather reports can be used to further corroborate the time a photo or video was taken, generally by showing that the weather depicted in the imagery coincided with the reported weather on the presumed date. However, historical weather analysis can be unreliable, and the weather conditions in photos or videos can now be altered with the availability of algorithmic and deepfake technology.<sup>42</sup>

#### Reverse image and video search

Reverse image search may also be used for chronolocation. Reverse image search may be useful for testing hypotheses about when a photo or video was created. For instance, if a caption for a photo alleges it was taken on a specific date (e.g. December 2017), but a reverse image search shows that the image existed online earlier (e.g. February 2014), this would indicate that the photo was not actually taken on the alleged date. However, if the reverse image search yields no results, this does not mean that the content did not exist previously, as search databases only account for a small percentage of online information. In general, reverse image searches can provide useful information to test investigators' hypotheses when they

<sup>42</sup> Samantha Cole, 'Watch an Algorithm Turn Winter Into Summer in Any Video' (Vice, 5 December 2017) <a href="https://www.vice.com/en/article/xwvz9a/watch-an-algorithm-turn-winter-into-summer-in-any-video-image-translation">https://www.vice.com/en/article/xwvz9a/watch-an-algorithm-turn-winter-into-summer-in-any-video-image-translation</a>>.

yield results, but if there are no results, it should not be assumed that the photo or video did not exist previously.

#### Satellite imagery comparison

Satellite imagery comparisons can be used for chronolocation when there are changes to an area over time that are visible on satellite imagery. For instance, historical satellite imagery can be used to narrow down a timeframe for when buildings were constructed or destroyed.<sup>43</sup> By viewing imagery from different time periods, analysts can see when new details appear or disappear. For example, satellite imagery may indicate when buildings were burned down or cultural monuments desecrated,<sup>44</sup> or may show disturbed earth, potentially indicating the presence of a mass grave. Importantly, however, satellite imagery (especially that which is publicly available, such as from Google Earth Pro) typically does not have clear historical imagery for each specific date and usually will only allow for a narrowed-down time frame of a few months.



**Figure H:** Screenshot of the digital platform developed by SITU Research for the International Criminal Court depicting the destruction of cultural heritage sites in Mali. The images show the site of the El Kounti mausoleum before and after destruction. (http://icc-mali.situplatform.com/).

**Key Takeaways:** There are multiple methods that investigators can use to provide an estimate of when a digital image was taken, including (but not limited to) metadata analysis, shadow analysis, image leads, historical weather analysis, upload time and date stamp analysis, reverse image and video searches, and satellite imagery comparisons. All have limitations. Reliable assessments for when a digital image was taken will consider and use multiple methods of analysis.

<sup>43</sup> Sam Dubberley and Joe Freeman, 'Killings, corruption, land grabs: human rights violations against the Rohingya today' (*Amnesty International*, 25 August 2020) <a href="https://citizenevidence.org/2020/08/25/rohingya-verification/">https://citizenevidence.org/2020/08/25/rohingya-verification/</a>>.

<sup>44</sup> Benjamin Strick, 'Geolocation of Infrastructure Destruction in Cameroon: A Case Study of Kumbo and Kumfutu' (Bellingcat, 21 November 2018) <a href="https://www.bellingcat.com/resources/case-studies/2018/11/21/geolocation-infrastructure-destruction-cameroon-case-study-kumbo-kumfutu/">https://situ.nyc/research/projects/icc-digital-platform-timbuktu-mali</a>.
Timbuktu, Mali' <a href="https://situ.nyc/research/projects/icc-digital-platform-timbuktu-mali">https://situ.nyc/research/projects/icc-digital-platform-timbuktu-mali</a>.

## Conclusion

Digital open source imagery is increasingly used by courts, human rights treaty bodies, and other fact-finding bodies. It can constitute highly probative material for the forensic and judicial assessment of suspected violations of international human rights law, international humanitarian law and international criminal law and can be introduced by both prosecution and defence in criminal trials to support their case. However, the growing prevalence of open source photos and videos as evidence imports risks of misinterpretation or misplaced reliance, either on open source materials, or the investigators' analysis. The techniques and descriptions contained in this guide are designed to assist judges and fact-finders when evaluating digital open source materials. At the same time, this type of material will typically form part of a larger body of evidence before the court or fact-finding body. Digital open source information should ultimately be assessed according to the same overarching evidentiary rules that are generally applied by the particular institution or court, and subject to the institution or court's established burdens and standards of proof.

## Glossary

**Artificial intelligence (AI):** a branch of computer science dedicated to developing programming for machines to learn how to react to unknown variables and adapt to new environments.

**Chronolocation**: the corroboration of the dates and times of the events depicted in a piece of information, usually visual imagery. For example, it may be possible to determine the time of day a photograph was taken by examining the length of the shadows made by sunlight, along with other indicators.

**Cryptographic hash value**: calculations that can be run on any type of digital file to generate a fixed-length alphanumeric string that can be used as evidence that a digital file has not been modified since that content was hashed. This string will remain the same every time the calculation is run as long as the file does not change.

**Dark web**: the part of the Internet that is only accessible by means of specialized software, and allows users and website operators to remain anonymous and untraceable.

**Deep web**: the part of the Internet that is not indexed and therefore is not accessible via search engines.

**Digital open source information**: publicly available information in digital format, which is generally acquired from the Internet.

**Geolocation**: the identification or estimation of the location of an object or an activity, or the location from which an item was generated. For example, it may be possible to determine the location from which a video or photograph downloaded from the Internet was taken using geolocation techniques. Such techniques could include, for example, identifying unique geographic features in a photograph with their actual location on a map.

**Metadata**: are data about data. They contain information about an electronic file that is either embedded in or associated with the file. Metadata often include a file's characteristics and history, such as its name, size, and dates of creation and modification. Metadata may describe how, when, and by whom or what a digital file was collected, created, accessed, modified and formatted.

**Open source information**: information that any member of the public can observe, purchase or request, without requiring special legal status or unauthorized access.

**Pseudonymization:** the processing of personal data in such a manner that the information can no longer be attributed to a specific data subject without the use of additional information.

Reverse image/video search: A reverse search involves uploading an image or video to a search engine so that the search algorithm can identify other copies of the same or similar images on the internet. The limitation of a reverse image search is that it only scans within a search engine's database, which includes a small percentage of the content currently on the internet. It does not, for example, include materials on the <a href="deep web">deep web</a> (which is not indexed to search engines such as Google) or the <a href="dark web">dark web</a> (the part of the Internet that can only be accessed through specialized software, such as the Tor browser).

**Synthetic media:** also referred to as generative media, is defined as visual, auditory, or multimodal content that has been generated or modified by algorithm (commonly via <u>artificial intelligence</u>). Such outputs are often highly realistic, would not be identifiable as synthetic to the average person, and may simulate artifacts, persons, or events.

**Verification:** refers to the process of establishing the accuracy or validity of information that has been collected online. The source, the content, and the digital item or file should be considered collectively and compared for consistency.

## Acknowledgements

Translation and production was funded by Swansea University's ESRC Impact Acceleration Account and the TRUE project at Swansea University, funded by UKRI Frontier Research Grant EP/X016021/1. The work was also supported by the Institute for Humanities and Social Sciences (IHSS) at Queen Mary University of London.

We wish to thank the following individuals, all of whom proved insightful comments on various aspects of the text at different stages of the drafting process: Dato' Shyamala Alagendra, international criminal lawyer; Hadi Al Khatib, Managing Director, Mnemonic; Siobhán Allen, Senior Lawyer, Global Legal Action Network (GLAN); shirin anlen, Media Technologist, Pavlo Bogachenko, Senior Associate, DLA Piper; Her Excellency Judge Solomy Bossa, Judge, Appeals Chamber, International Criminal Court; Jacobo Castellanos, Coordinator, Camille Chabot, Researcher, Berkeley School of Law Human Rights Center and Peking University; Her Excellency Judge Margaret De Guzman, Judge International Residual Mechanism for Criminal Tribunals; Dr Jeff Deutch, Senior Researcher Mnemonic; Sam Dubberley; Michael Elsanadi, Open Source Investigator, Mnemonic; Jessica Gavron, Legal Director, European Human Rights Advocacy Centre; Dr Matthew Gillett, Senior Lecturer, School of Law and Human Rights, University of Essex; Jonathan Hak KC; Anne Hausknecht, PhD student, TRUE Project, Swansea University; Peter Haynes KC; Professor Laurence R. Helfer, Professor of Law, Duke University, Member of the UN Human Rights Committee; Gabriele Juodkaite-Granskiene, Judge, Supreme Court of Lithuania; Koen Kluissen, Detective Inspector and Open Source Investigator (International Crimes), Netherlands Public Prosecution Service; Steve Kostas, Senior Legal Officer, Open Society Justice Initiative; Her Excellency Judge Joanna Korner, Judge International Criminal Court; Professor Philip Leach, Professor of Human Rights Law, Middlesex University London; Kateryna Latysh, MSCA4Ukraine Postdoctoral Fellow, Vilnius University and Associate Professor, Yaroslav Mudryi National Law University; Nema Milaninia, Special Advisor to the US Ambassador-at-Large for Global Criminal Justice; Dearbhla Minogue, Senior Lawyer, Global Legal Action Network (GLAN); Judy Mionki, Defence Counsel, International Criminal Court, Africa Region Liaison Officer, International Bar Association's Human Rights Law Committee; Yvonne Ng, Archives Program Manager, WITNESS; Raluca Racusan; His Honour Judge Keith Raynor, Judge, England and Wales; Yaroslavna Sychenkova, independent consultant; Professor Yuval Shany, Hersch Lauterpacht Chair in International Law, Hebrew University of Jerusalem; Konstantina Stavrou, PhD Candidate, University of Vienna; Benjamin Strick, Centre for Information Resilience; Hryhorii Zhurakivskyi, European Union Assistance Mission to Ukraine.

Thank you also to all the judges and former judges who provided comments but wished to remain anonymous.



Evaluating digital open source imagery:
A guide for judges and fact-finders