Systematic Protocol for Open-Source Intelligence Gathering

Open-source intelligence (OSINT) collection requires discipline, structure, and legal awareness. Ad-hoc searching wastes time and creates gaps. This checklist guides you through a systematic OSINT collection process — from scoping requirements through to documentation and source validation.

Phase 1 — Define Collection Requirements

Before You Search

- ■ The intelligence requirement or question is clearly defined in writing.
- ■ The subject type is identified: person / organisation / location / event / topic.
- ■ The time period of interest is specified (historical, current, ongoing).
- ■ Geographic scope is defined (local, national, international).
- ■ The minimum acceptable level of source reliability is established.
- ■ Legal authority or authorisation to collect is confirmed.
- ■ Privacy and data protection obligations have been reviewed.

Phase 2 — Search Strategy

Search Term Preparation

- ■ Primary search terms (name, alias, username, entity) are listed.
- ■ Alternative spellings, transliterations, and abbreviations are documented.
- ■ Boolean operators (AND, OR, NOT) are planned for key searches.
- ■ Date range filters are set where recency matters.
- ■ Language variants are considered for international subjects.

Search Engine Coverage

- ■ Google (standard and verbatim search) completed.
- ■ Bing completed (often indexes different content to Google).
- ■ DuckDuckGo completed (no personalisation bias).
- ■ Google advanced search operators applied (site:, filetype:, intitle:, inurl:).
- ■ Cached and archived versions checked where live pages are unavailable.

Phase 3 — Source Categories

Social Media

- ■ LinkedIn — professional history, connections, endorsements.
- ■ Facebook — personal network, groups, check-ins, life events.
- ■ X/Twitter — posts, retweets, follower/following lists, bio history.
- ■ Instagram — location tags, tagged photos, story archives.
- ■ TikTok / YouTube — video content, channel history, comments.
- ■ Forums and community boards (Reddit, specialist forums) searched.
- ■ Username consistency checked across platforms (use Sherlock / Namechk).

Public Records & Official Sources

- ■ Company registrations / ABN / ASIC / equivalent national registry.
- ■ Court records and legal proceedings checked.
- ■ Electoral roll / voter registration (where publicly accessible).
- ■ Government gazette and procurement portals searched.
- ■ Property ownership records searched (land titles / council records).
- ■ Charity and NGO registrations reviewed.
- ■ Professional licensing and disciplinary registers checked.

News & Media

- ■ National and local news archives searched.
- ■ Trade and industry publications searched.
- ■ Press release databases searched (PR Newswire, Business Wire).
- ■ Broadcast media transcripts searched where available.
- ■ Google News alerts set for ongoing monitoring.

Domain, IP & Technical Sources

- ■ WHOIS domain registration records checked.
- ■ Reverse IP lookup completed.
- ■ Shodan / Censys search completed for infrastructure exposure.
- ■ SSL certificate transparency logs searched (crt.sh).
- ■ Wayback Machine / archive.org used for historical snapshots.
- ■ Email header analysis completed where applicable.

Document & File Sources

- ■ Google dorking for file types: filetype:pdf, filetype:xlsx, filetype:docx.
- ■ Scribd / SlideShare / Academia.edu searched.
- ■ Government and organisational document repositories searched.
- ■ Metadata extracted and reviewed from obtained documents.
- ■ Leaked credential databases checked (HaveIBeenPwned).

Geospatial & Location Sources

- ■ Google Maps / Street View reviewed for physical location.
- ■ Satellite imagery reviewed (Google Earth, Sentinel Hub).
- ■ Geotagged social media posts searched.
- ■ Flight tracking (FlightAware, FlightRadar24) checked if relevant.
- ■ Vessel tracking (MarineTraffic) checked if relevant.

Phase 4 — Source Evaluation

Apply the following criteria to every source before including it in your assessment.

Criterion	Questions to Ask	Rating
Reliability	Is the source known and consistent? Does it have a track record of accuracy?	A B C D E F
Credibility	Is this specific piece of information corroborated by other sources?	1 2 3 4 5 6
Currency	Is the information recent enough to be relevant to the current requirement?	■ Current ■ Dated
Objectivity	Does the source have a known bias, agenda, or conflict of interest?	■ Neutral ■ Biased
Provenance	Can the original source be traced? Is this first-hand or second-hand reporting?	■ Primary ■ Secondary

Reliability scale: *A=Completely reliable, B=Usually reliable, C=Fairly reliable, D=Not usually reliable, E=Unreliable, F=Cannot be judged.* ***Credibility scale:*** *1=Confirmed, 2=Probably true, 3=Possibly true, 4=Doubtfully true, 5=Improbable, 6=Cannot be judged.*

Phase 5 — Documentation & Legal Compliance

Recording & Attribution

- ■ Every piece of collected information has a source URL and date of access recorded.
- ■ Screenshots taken of key web content (pages can be deleted or modified).
- ■ File metadata preserved and recorded.
- ■ Collection log maintained with timestamps.
- ■ Chain of custody documented for any obtained files or data.

Legal & Ethical Compliance

- ■ No collection methods used that require authentication without authorisation.
- ■ No scraping or automated tools used in violation of terms of service.
- ■ Personal information collected is necessary and proportionate to the requirement.
- ■ Collected data is stored securely and access is limited to those with a need to know.
- ■ Retention and disposal obligations have been identified.
- ■ Collection activity has not interfered with any active law enforcement operation.

Need a Different Format?

• A Word (.docx) editable version is available on request.
• Email: info@theintelanalystacademy.com.au — Subject: 'OSINT Checklist — Format Request'