INTELLIGENCE · OPEN-SOURCE INTELLIGENCE

"Open-Source Intelligence (OSINT): Methods and Sources"

Open-source intelligence turns publicly available information into knowledge that supports decisions. This article explains what counts as an open source, the main families of sources, a practical workflow from question to report, and the limits and ethics of the discipline.

ModQu research article · 1,939 words · 20 references (APA 7th edition)

Most of what any organisation needs to know about its competitors, markets, suppliers and risks is not secret. It is published in newspapers, company filings, court records, patents, academic papers, websites and social media posts, or visible in satellite images and public data sets. Open-source intelligence, usually shortened to OSINT, is the discipline of finding, collecting, checking and analysing this publicly available information to answer a specific question. This article explains what OSINT is, the main families of sources, a practical workflow and the limits and ethical rules that apply.

What OSINT is

The word "open" refers to the source, not to the result. OSINT uses information that anyone can lawfully obtain, by reading, buying, requesting or observing it. The intelligence comes from what is done with that information. A widely used definition, set out in American law and discussed by Best and Cumming (2007), describes OSINT as intelligence produced from publicly available information that is collected, exploited and disseminated in a timely manner to an appropriate audience for the purpose of addressing a specific intelligence requirement. The final clause matters: collecting interesting material is not OSINT unless it answers a question someone needs answered.

Open sources have always been part of intelligence. During the Second World War, governments monitored foreign radio broadcasts and newspapers, and those monitoring services continued throughout the Cold War. Mercado (2004) describes how much useful knowledge such work produced, often at a fraction of the cost of secret collection. Steele (2007), a long-standing advocate, argued that open sources could meet most of an organisation's information needs if they were exploited systematically.

The internet changed the scale of the discipline. Glassman and Kang (2012) describe how the growth of online publishing and social media turned OSINT from a supporting activity into a central one, and how it increasingly involves collaboration among many people outside traditional intelligence organisations. After reviewing intelligence failures before the Iraq war, the Commission on the Intelligence Capabilities of the United States Regarding Weapons of Mass Destruction (2005) recommended that open sources be given a dedicated organisation and much greater priority. Lowenthal (2022) notes that open sources are often the first place an analyst should look, because they are fast, cheap and can be shared without revealing secret methods.

Williams and Blum (2018), in a study for the RAND Corporation, describe a "second generation" of OSINT, driven by the sheer volume and variety of online information, new tools for collecting and processing it, and the fact that commercial firms, journalists and private citizens now practise it alongside governments.

Families of open sources

Open sources can be grouped into several families, each with typical strengths and weaknesses.

Families of open sources

What each family offers and what to watch for

Family	Examples	Strengths	Cautions
Media	News, trade press, broadcasts	Timely, wide coverage, local detail	Errors, slant, copying between outlets
Internet and social media	Posts, forums, video, reviews	Eyewitness content, sentiment, speed	Unrepresentative, fakes, bots
Public records	Company filings, courts, patents, procurement	Authoritative, often legally required	Delays, gaps, varied formats
Grey and academic literature	Research papers, think-tank and industry reports	Depth, expert analysis	Slow; agendas; uneven quality
Commercial data	Databases, market reports, data vendors	Structured, comparable	Cost; opaque methods
Geospatial	Satellite and aerial imagery, maps	Independent physical evidence	Needs expert interpretation

Compiled from Steele (2007), Williams and Blum (2018) and Pastor-Galindo et al. (2020)

ModQu.com

Figure 1. Families of open sources. Each family has typical strengths and typical weaknesses; good OSINT combines several. Compiled from Steele (2007), Williams and Blum (2018) and Pastor-Galindo et al. (2020); table by ModQu.

Media. Newspapers, broadcasters, trade publications and specialist newsletters report events quickly and often include local detail. But outlets make errors, have editorial positions and frequently repeat each other, so ten articles may rest on a single original source.

Internet and social media. Posts, videos, forums, reviews and messaging channels provide eyewitness material and a window on opinion. Omand et al. (2012) coined the term social media intelligence (SOCMINT) and argued that it needs its own methods and safeguards. Social media users are not representative of the population, and false material spreads easily. Vosoughi et al. (2018) found that false news stories on a major platform spread further and faster than true ones.

Public records. Company registrations and filings, court records, land registries, patents, regulatory decisions, procurement notices and trade statistics are often authoritative because they are legally required. They can be slow to appear, incomplete and published in inconsistent formats across jurisdictions.

Grey and academic literature. Research papers, conference presentations, theses, think-tank studies and industry association reports provide depth and expert analysis. They are slower than news and may reflect the agenda of their sponsors.

Commercial data. Paid databases, market research reports and data vendors provide structured information that is easy to compare. Their methods are not always transparent.

Geospatial sources. Satellite and aerial imagery, maps and data contributed by the public allow independent observation of physical activity. Donaldson and Storeygard (2016) describe how economists use satellite data to measure activity where official statistics are weak. Goodchild (2007) described how members of the public act as "sensors", contributing geographic information that once only official agencies collected.

Pastor-Galindo et al. (2020) survey the wide range of tools that now support OSINT, from search engines and archive services to tools for analysing images, networks and metadata, and note that their effective use depends on trained analysts.

A practical workflow

OSINT work follows a sequence that resembles the wider intelligence process, with particular emphasis on preserving and verifying material.

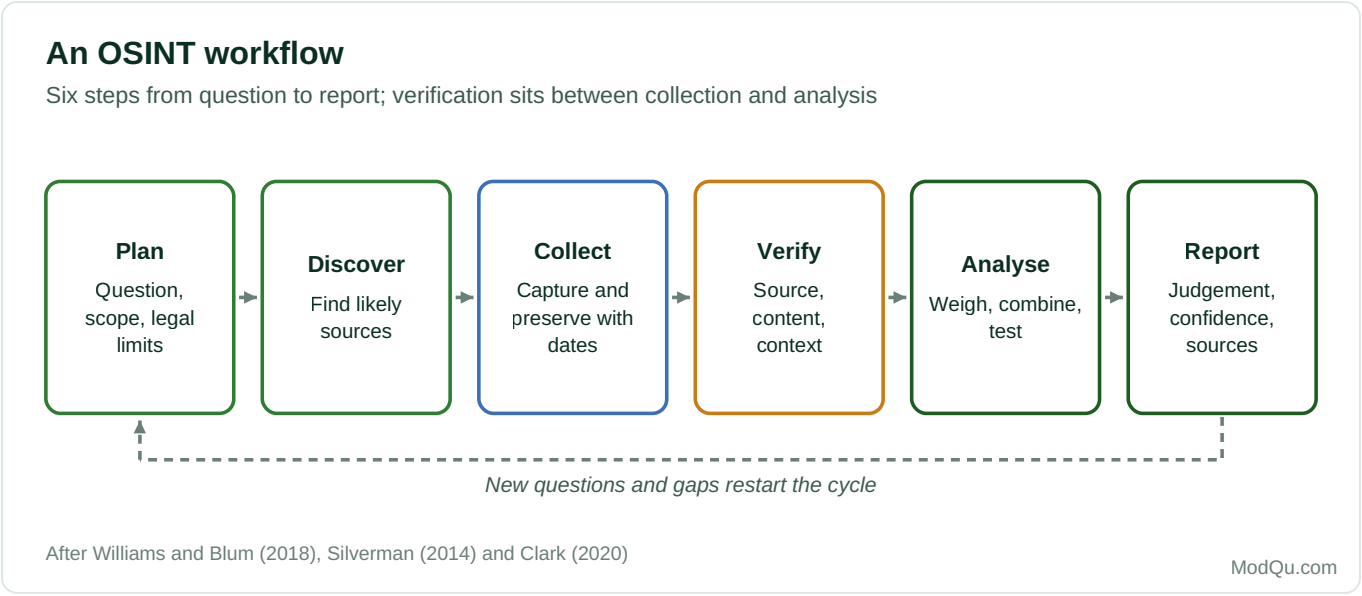


Figure 2. An OSINT workflow. Verification sits between collection and analysis, and each report usually raises new questions. After Williams and Blum (2018), Silverman (2014) and Clark (2020); diagram by ModQu.

1. **Plan.** Define the question, the decision it supports, the time available and the legal and ethical limits. Clark (2020) recommends building a model of the target, the organisation, market or problem being studied, and using it to identify what information is needed.
2. **Discover.** Identify where relevant information is likely to be found: which registries, publications, platforms, databases and experts. Good analysts keep lists of reliable sources for each type of question.
3. **Collect and preserve.** Capture material with its source, date and address. Online content changes and disappears; screenshots, saved copies and archive services make it possible to show later what was found and when.
4. **Verify.** Check who created each item, whether it is authentic and whether it means what it appears to mean. Silverman (2014), in a handbook for journalists, sets out practical checks on the source, the content, the date and the location of online material.
5. **Analyse.** Combine items, weigh their reliability and test alternative explanations. Heuer (1999) and Pherson and Heuer (2021) describe structured techniques, such as listing competing hypotheses and checking which evidence is consistent with each, that guard against seizing on the first plausible story.
6. **Report.** State the judgement, the confidence it deserves and the sources on which it rests, so that readers can check them.

Higgins (2021) describes how the investigative group Bellingcat applied this kind of workflow to establish facts about conflicts and crimes using only open sources: matching landscape features in videos with satellite images to find where they were filmed, using shadows to estimate the time of day, and tracing the online histories of individuals and units. The methods depend less on special access than on patience, careful checking and transparency about how each conclusion was reached.

Limits of OSINT

Open sources have real weaknesses. Hulnick (2002) warned of several downsides: the volume of material can overwhelm analysts; much of it is unreliable; filtering and checking it takes considerable effort; and organisations may over-rely on what is easily available while neglecting what is not. Some questions simply cannot be answered from public information, because the relevant facts are deliberately hidden.

Open sources can also be manipulated. A competitor, a government or an interest group may publish misleading material precisely because it knows others are watching. Fake accounts, fabricated documents and edited images are increasingly easy to produce. The more an organisation relies on open sources, the more important verification becomes.

Finally, what is visible online is shaped by who publishes and by the platforms' own design. Absence of information is not evidence of absence: a firm with no online presence in a region may still be active there.

Ethics and law

That information is public does not mean every use of it is acceptable. Zimmer (2010), analysing a research project that collected students' social media profiles, showed that people can have reasonable expectations about how their publicly visible information will be used, and that combining scattered public details can reveal more than any one source intended. In Canada, the Personal Information Protection and Electronic Documents Act (2000) governs how private organisations collect, use and disclose personal information in commercial activity, including information gathered online. Platforms' terms of service, copyright and laws against unauthorised access to computer systems also apply.

Good practice for OSINT in business includes: collecting only what the question requires; focusing on organisations and public roles rather than private lives; never misrepresenting identity to obtain information; respecting terms of use and technical restrictions; and recording sources so that conclusions can be checked. Omand et al. (2012) argue that public trust depends on intelligence collection being necessary, proportionate and properly authorised, principles that translate naturally to private organisations.

Skills and organisation

OSINT is often described as a set of tools, but its quality depends on people. Effective analysts combine several skills: knowing where particular kinds of information are published in each country and industry; searching precisely, with careful choice of terms, languages and filters; reading critically, with attention to who wrote something and why; handling basic data and images; and writing clear judgements with stated confidence. Language skills and local knowledge are especially valuable, because much relevant material is never translated.

Organisations can support this work in simple ways. A shared register of trusted sources for recurring questions saves time and improves consistency. A standard format for recording what was found, where and when makes work reusable and checkable. Separate, clearly managed accounts and devices for research protect the organisation's own systems and avoid accidental disclosure of its interest. And a short legal and ethical checklist, agreed in advance, removes uncertainty about what analysts may and may not do.

Common mistakes

- **Starting without a question.** Collection without a defined requirement produces large files and few answers.
- **Counting repetitions as confirmation.** Many reports may trace back to one original claim.
- **Failing to preserve material.** A web page that changes or disappears cannot be cited or rechecked.
- **Treating visibility as reality.** What is easy to find online may not reflect what matters most.
- **Hiding uncertainty.** Readers need to know which judgements rest on strong evidence and which on inference.

An illustration

A Canadian manufacturer of industrial pumps hears a rumour that a foreign competitor plans to open a plant in Ontario. An analyst begins by defining the question: is the competitor building a plant, where, of what size and on what timetable? Searching provincial corporate registries reveals a newly registered subsidiary. Municipal council minutes show a zoning application for an industrial site under that subsidiary's name. The competitor's job postings on its own website and on employment platforms list positions for a plant manager and production engineers in the region. Satellite imagery, compared across several months, shows ground clearing at the site. A trade publication quotes a regional economic development official describing "a new advanced manufacturing investment".

No single source is conclusive, but together they support the judgement, with high confidence, that the competitor is building a plant, and with moderate confidence that it will open within eighteen months. The analyst records every source, the date it was checked and the reasoning, and identifies indicators to watch: building permits, equipment import records and further hiring.

Conclusion

OSINT turns publicly available information into intelligence by applying the same discipline as any other intelligence work: a clear question, systematic collection, careful verification, structured analysis and transparent reporting. Its sources are vast, inexpensive and increasingly rich, but also noisy, uneven and open to manipulation. Organisations that treat open sources as a craft, with trained analysts, recorded methods and clear ethical rules, can answer most of the questions they face without any secret access at all.

References

Best, R. A., Jr., & Cumming, A. (2007). *Open source intelligence (OSINT): Issues for Congress* (CRS Report RL34270). Congressional Research Service.

Clark, R. M. (2020). *Intelligence analysis: A target-centric approach* (6th ed.). CQ Press.

Commission on the Intelligence Capabilities of the United States Regarding Weapons of Mass Destruction. (2005). *Report to the President of the United States*. U.S. Government Printing Office.

Donaldson, D., & Storeygard, A. (2016). The view from above: Applications of satellite data in economics. *Journal of Economic Perspectives*, *30*(4), 171–198. https://doi.org/10.1257/jep.30.4.171

Glassman, M., & Kang, M. J. (2012). Intelligence in the internet age: The emergence and evolution of Open Source Intelligence (OSINT). *Computers in Human Behavior, 28*(2), 673–682. https://doi.org/10.1016/j.chb.2011.11.014

Goodchild, M. F. (2007). Citizens as sensors: The world of volunteered geography. *GeoJournal, 69*(4), 211–221. https://doi.org/10.1007/s10708-007-9111-y

Heuer, R. J., Jr. (1999). *Psychology of intelligence analysis*. Center for the Study of Intelligence, Central Intelligence Agency.

Higgins, E. (2021). *We are Bellingcat: An intelligence agency for the people*. Bloomsbury.

Hulnick, A. S. (2002). The downside of open source intelligence. *International Journal of Intelligence and CounterIntelligence, 15*(4), 565–579. https://doi.org/10.1080/08850600290101767

Lowenthal, M. M. (2022). *Intelligence: From secrets to policy* (9th ed.). CQ Press.

Mercado, S. C. (2004). Sailing the sea of OSINT in the information age. *Studies in Intelligence, 48*(3), 45–55.

Omand, D., Bartlett, J., & Miller, C. (2012). Introducing social media intelligence (SOCMINT). *Intelligence and National Security, 27*(6), 801–823. https://doi.org/10.1080/02684527.2012.716965

Pastor-Galindo, J., Nespoli, P., Gómez Mármol, F., & Martínez Pérez, G. (2020). The not yet exploited goldmine of OSINT: Opportunities, open challenges and future trends. *IEEE Access, 8*, 10282–10304. https://doi.org/10.1109/ACCESS.2020.2965257

Personal Information Protection and Electronic Documents Act, S.C. 2000, c. 5 (Canada).

Pherson, R. H., & Heuer, R. J., Jr. (2021). *Structured analytic techniques for intelligence analysis* (3rd ed.). CQ Press.

Silverman, C. (Ed.). (2014). *Verification handbook: A definitive guide to verifying digital content for emergency coverage*. European Journalism Centre.

Steele, R. D. (2007). Open source intelligence. In L. K. Johnson (Ed.), *Handbook of intelligence studies* (pp. 129–147). Routledge.

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science, 359*(6380), 1146–1151. https://doi.org/10.1126/science.aap9559

Williams, H. J., & Blum, I. (2018). *Defining second generation open source intelligence (OSINT) for the defense enterprise* (Report RR-1964-OSD). RAND Corporation.

Zimmer, M. (2010). "But the data is already public": On the ethics of research in Facebook. *Ethics and Information Technology, 12*(4), 313–325. https://doi.org/10.1007/s10676-010-9227-5