INTELLIGENCE · OPEN-SOURCE INTELLIGENCE

"OSINT for Business: Public Records, Filings and Media"

Public records, regulatory filings and the media are the backbone of business open-source intelligence. This article explains what each source reveals, how to read filings and news systematically, how text analysis can help, and the pitfalls and rules that apply.

ModQu research article · 2,637 words · 24 references (APA 7th edition)

When a business wants to understand a competitor, a supplier, a customer or a potential partner, the most reliable information often comes from sources that the subject was required to publish. Companies must register with governments, file financial reports, apply for permits and patents, and defend themselves in court. Journalists report on their activities. Each of these leaves a public record. Business open-source intelligence makes systematic use of these records, together with media coverage, to build an accurate picture of organisations and markets.

This article explains what the main categories of public records and filings reveal, how to read them, how text analysis is changing the work, what the media contribute and what pitfalls and rules to keep in mind.

Why public records matter

Public records have a particular advantage over other open sources: many are produced under legal obligation, often with penalties for false statements. A company's audited financial statements, its registration documents and the decisions of courts and regulators are not marketing material. They may be incomplete or late, but they are generally more reliable than press releases or social media.

Kahaner (1996), in one of the first practical guides to competitive intelligence, emphasised that much of what firms need to know about competitors can be found in such documents if one knows where to look. Steele (2007) argued more broadly that most intelligence needs can be met from open sources when they are exploited systematically. Calof and Wright (2008) note that competitive intelligence as a professional field grew up around exactly this kind of legal, ethical collection from public information.

The main categories of public records

What public records reveal about a business

Common record types, what they show, and where to find them in Canada

Record	What it can reveal	Canadian examples
Corporate registrations	Legal entities, directors, addresses, new subsidiaries	Corporations Canada; provincial registries
Securities filings	Results, strategy, risks, executive pay, major contracts	SEDAR+ (Canada); EDGAR (United States)
Court and tribunal records	Disputes, insolvency, regulatory actions	Court registries; CanLII decisions
Patents and trademarks	Technology direction, new brands, launch timing	Canadian Intellectual Property Office
Government contracts	Customers, prices, partners, capabilities	Federal and provincial tender portals
Permits and zoning	New sites, expansions, closures	Municipal council minutes and permits
Lobbying registries	Policy interests and the officials contacted	Federal and provincial lobbyist registries
Trade and customs data	Import and export flows by product and country	Statistics Canada trade data

Compiled by ModQu; record types after Kahaner (1996) and Fleisher and Bensoussan (2015)

ModQu.com

Figure 1. What public records reveal about a business. Each record type answers different questions; together they can show structure, strategy, problems and plans. Compiled by ModQu; record types after Kahaner (1996) and Fleisher and Bensoussan (2015).

Corporate registrations

Every incorporated business must register with a federal or provincial authority. Registrations show legal names, registered addresses, directors and, in some jurisdictions, officers and shareholders. They reveal corporate structure: which subsidiaries a group owns, when new entities were created and where. A newly registered subsidiary in a new province or country can be an early sign of expansion.

Securities filings

Publicly listed companies must file annual and quarterly reports, management's discussion and analysis, information circulars, material change reports and other documents. In Canada these are available through SEDAR+; in the United States, through the Securities and Exchange Commission's EDGAR system. Healy and Palepu (2001) review the extensive research on corporate disclosure, which shows that firms disclose information both because they must and because doing so lowers their cost of capital. Filings typically contain:

- financial results by segment and geography;
- descriptions of the business, its strategy and its competitors;
- discussions of risks that management considers material;
- information on major customers, suppliers and contracts;
- executive compensation and board composition.

Private companies file far less, but they may still appear in records of the companies they deal with, in court files and in government contracts.

Court and regulatory records

Lawsuits, insolvency proceedings, regulatory decisions and tribunal rulings reveal disputes with customers, suppliers, employees and competitors. They can expose problems that firms would rather not publicise, such as product defects or contract failures. Many Canadian court and tribunal decisions are published through free legal databases, while dockets and filed documents may require a request to the court registry. Insolvency filings deserve particular attention, because the creditor lists attached to them can reveal a firm's suppliers, customers and the amounts owed to each.

Patents and trademarks

Patent applications reveal where firms are investing in technology, often years before products appear. Griliches (1990) showed that patent counts, despite their limitations, are useful indicators of inventive activity, and Ernst (2003) described how patent information can support competitive analysis and technology planning. Trademark applications can signal new brands and product launches.

Government contracts, permits and lobbying

Public procurement notices and contract awards show who sells what to government, at what price and with which partners. Municipal permits and zoning applications reveal new sites, expansions and closures. Lobbyist registries show which policy issues an organisation is pursuing and which officials it has contacted.

Trade data

National statistics agencies publish import and export data by product and partner country. Although individual firms are not named, detailed product codes can reveal market trends, new sources of supply and the size of foreign competition.

Reading filings systematically

Filings are long, and much of their content is repeated from year to year. Reading them well means knowing where change is most informative and comparing documents across time and across firms.

Useful practices include:

1. **Compare year on year.** Changes in segment reporting, the description of the business or the list of competitors often signal shifts in strategy.
2. **Read the risk factors.** Firms must describe risks they consider material. New risks, or risks that move up the list, indicate what management is worried about.
3. **Check the notes.** Commitments, contingencies, related-party transactions and subsequent events are often in the notes to the financial statements rather than the main text.
4. **Follow the money.** Capital spending by segment, research and development budgets and acquisitions show where firms are placing their bets.
5. **Compare peers.** Reading competitors' filings side by side reveals differences in margins, growth and stated priorities.

Cohen et al. (2020) showed that changes in the language of annual and quarterly reports carry information that investors are slow to notice. Firms whose reports changed substantially from one year to the next, particularly in their discussion of risks, tended to perform worse afterwards. The authors called the tendency of investors to overlook such changes "lazy prices". For business intelligence, the lesson is direct: the parts of a document that change are often the parts worth reading.

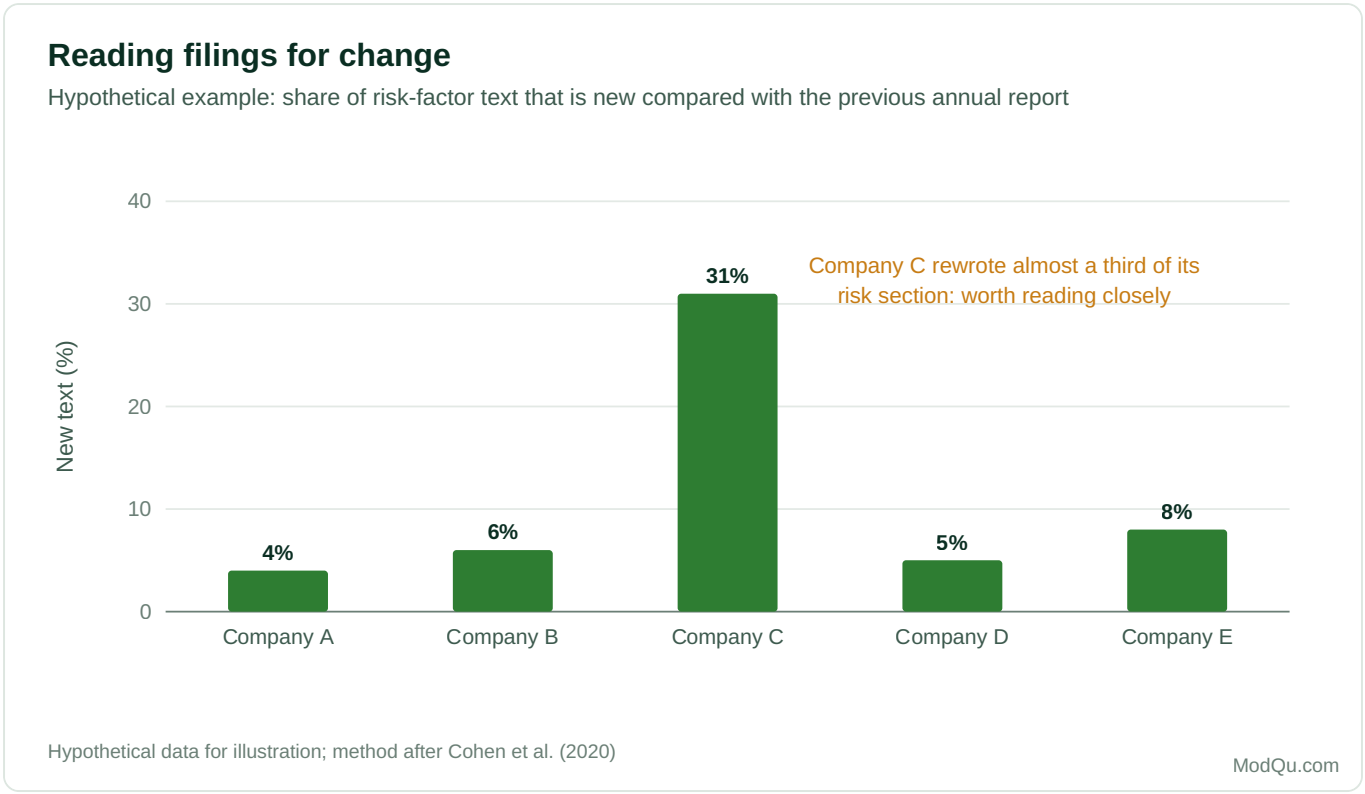


Figure 2. Reading filings for change. A simple comparison of text between years flags the documents worth reading closely. Hypothetical data for illustration; method after Cohen et al. (2020); chart by ModQu.

Hoberg and Phillips (2016) used the product descriptions in firms' annual reports to measure how similar firms are to each other, producing industry groupings that change over time as firms move into new markets. Their approach shows that filings can be used to identify competitors that conventional industry classifications miss.

Text analysis of filings and news

The volume of filings and news is too great for anyone to read in full. Text analysis offers ways to screen, classify and measure documents at scale. Gentzkow et al. (2019) and Grimmer and Stewart (2013) review methods for treating text as data, from counting words in predefined dictionaries to statistical models that learn topics and classify documents.

Two lessons from this research matter most for practitioners. The first is that context matters. Loughran and McDonald (2011) found that general-purpose word lists misclassify much of the language in financial documents: words such as "liability", "tax" and "cost" are counted as negative in everyday usage but are neutral in a financial report. They built word lists specific to finance, which performed better. Li (2010) found that a machine-learning approach trained on forward-looking statements in filings predicted future performance better than general dictionaries. In their later survey, Loughran and McDonald (2016) emphasise that results depend heavily on choices of method and that methods should be transparent and replicable.

The second lesson is that automated methods need validation. Grimmer and Stewart (2013) stress that all text models are simplifications and must be checked against human reading. A tool that flags documents for attention is useful; one whose classifications are accepted without checking can mislead.

What the media contribute

Journalism adds information that filings do not contain: interviews with employees, customers and former executives, coverage of local events, investigation of problems firms would prefer to hide, and analysis of trends. Kovach and Rosenstiel (2014) describe verification as the essence of journalism, distinguishing it from entertainment, propaganda and fiction; good reporting is therefore often a partly verified source in its own right.

Research shows that the press plays a real role in uncovering problems. Miller (2006) found that the press often reported on accounting fraud before it was officially announced, both by passing on analysis from other sources and by original investigation. Dyck et al. (2010), studying how corporate frauds in large American companies came to light, found that employees, the media and industry regulators were among the most important sources, more so than auditors and securities regulators.

Media coverage also affects how markets see firms. Tetlock (2007) found that pessimistic language in a widely read daily newspaper column predicted downward pressure on share prices, followed by a partial reversal. Fang and Peress (2009) found that shares of firms with no media coverage earned higher returns than those with heavy coverage, consistent with the idea that coverage spreads information and reduces uncertainty.

For intelligence purposes, the media are most useful when their limitations are kept in mind. Many articles repeat press releases. Several outlets may rely on the same wire report. Trade publications may depend on advertisers in the industry they cover. Local news, specialist publications and foreign-language sources often contain details that national outlets miss.

Alternative data

Beyond traditional records and media, a growing range of "alternative data" is used to measure business activity: prices collected from retailers' websites, satellite images of factories and car parks, shipping data, job postings and aggregated card transactions. Cavallo and Rigobon (2016) describe how online prices collected daily from many retailers can measure inflation and price behaviour. Zhu (2019) found that the availability of alternative data, such as satellite images of retailers' car parks and consumer transaction data, made share prices more informative and appeared to discipline managers' behaviour. These sources complement public records by offering timely measures of activity between official reports.

Combining sources

Each type of record answers part of a question. The strongest conclusions come from triangulation: finding the same fact, or consistent facts, in independent sources. A plan to build a new plant, for example, may appear first as a new subsidiary in a corporate registry, then as a zoning application, then as job postings, then in a filing that mentions capital

commitments, and finally in local news. When independent records agree, confidence rises. When they conflict, the conflict itself is informative: a firm may have announced an expansion to investors that it has not yet begun on the ground.

Triangulation also guards against deliberate deception. A press release can say anything; a building permit, a court filing and an audited statement are much harder to fabricate consistently.

Profiling a company: a checklist

When an organisation needs a quick but sound profile of another company, the following sequence covers the main public sources:

1. **Identity.** Confirm the exact legal name, registration number, jurisdiction and address. List parent companies and subsidiaries.
2. **People.** Identify directors and senior officers from registries and filings, noting recent changes.
3. **Finances.** For listed firms, extract revenue, margins, cash flow and debt by segment for at least three years. For private firms, look for credit reports, government contracts and filings by lenders.
4. **Strategy.** Read the business description, management's discussion and recent investor presentations for stated priorities.
5. **Risks and problems.** Read risk factors, notes on contingencies, court records and regulatory decisions.
6. **Activity.** Check patents, trademarks, permits, procurement awards and job postings for signs of new products, sites and capabilities.
7. **Reputation.** Review news coverage, trade press and, with care, customer reviews.
8. **Gaps.** Record what could not be found and what would be needed to fill the gap.

Each item should be recorded with its source and date, so that the profile can be updated rather than rebuilt.

Setting up ongoing monitoring

Most business intelligence needs are continuing rather than one-off. Instead of researching a competitor from scratch each time a question arises, organisations can monitor key sources routinely. Many registries, court systems and filing services allow alerts when new documents are filed under a given company name. News alerts can be set for company names, executives and products. Patent offices publish new applications regularly, and procurement portals list new awards.

A modest monitoring routine might assign each important competitor, customer and supplier to a named analyst, who reviews new filings, records and news weekly and writes a short note only when something changes. Over time, the notes form a history that makes it much easier to detect departures from normal patterns. The routine should also have a clear route for urgent findings: a court filing that threatens a key supplier, or a regulatory decision against a major customer, needs to reach the people who can act within hours, not at the next monthly meeting. Reviewing the monitoring list every six months, adding new competitors and dropping those that no longer matter, keeps the effort focused.

Pitfalls

Business OSINT has well-known pitfalls. Hulnick (2002) warned that the sheer volume of open information can overwhelm analysts and that reliability varies widely. Specific problems include:

- **Stale information.** Registries and filings may be months out of date. Always note the date of each record.
- **Name confusion.** Firms with similar names, or with complex groups of subsidiaries, are easily confused. Use registration numbers and addresses to confirm identity.
- **Selective disclosure.** Firms choose how to present information within legal requirements. A positive narrative in a filing does not cancel the risks in the notes.
- **Echoes in the media.** Repetition across many outlets is not independent confirmation.
- **Missing private companies.** Much of the economy consists of private firms that file little. Absence of records is not absence of activity.
- **Jurisdictional differences.** What must be disclosed varies between countries and provinces, so comparisons across borders need care.

Law and ethics

Public records are public, but their use is still subject to rules. Personal information in records, such as the home addresses of directors, is subject to privacy law; in Canada, the Personal Information Protection and Electronic Documents Act (2000) governs the collection, use and disclosure of personal information by private organisations in commercial activity. Some registries restrict bulk downloading or commercial reuse in their terms of use. Professional codes in competitive intelligence require practitioners to identify themselves honestly and to avoid misrepresentation when seeking information from people (Fleisher & Bensoussan, 2015).

An illustration

A Canadian distributor of medical supplies wants to understand a listed competitor that has been winning hospital contracts. An analyst reads the competitor's last three annual reports on SEDAR+ and notices that its risk factors now include a paragraph on dependence on a single overseas manufacturer, which did not appear before. Segment data show margins falling in the hospital business even as revenue grows. Provincial procurement portals show that the competitor won several large contracts at prices below the distributor's own bids. A court database reveals a dispute between the competitor and a former logistics provider over unpaid invoices. A regional newspaper reports that the competitor has laid off warehouse staff.

Together, the records suggest a competitor buying market share with low prices, dependent on one supplier and under financial strain. The distributor decides not to match the lowest prices but to emphasise reliability of supply in its next bids, and sets up monitoring of the competitor's filings and the court case. Gilad (2004) argues that this kind of early warning, detecting a rival's vulnerabilities and likely moves before they become obvious, is the main value of competitive intelligence.

Conclusion

Public records, regulatory filings and the media provide the backbone of business open-source intelligence. Their strength lies in legal obligation, independent reporting and the fact that they accumulate over time, allowing comparison. Using them well requires knowing where each kind of record is kept, reading documents for change, verifying identities and dates, using text analysis carefully and respecting privacy and terms of use. Analysts who master these sources can answer most questions about organisations and markets with evidence that others can check.

References

Calof, J. L., & Wright, S. (2008). Competitive intelligence: A practitioner, academic and inter-disciplinary perspective. *European Journal of Marketing, 42*(7/8), 717–730. https://doi.org/10.1108/03090560810877114

Cavallo, A., & Rigobon, R. (2016). The Billion Prices Project: Using online prices for measurement and research. *Journal of Economic Perspectives, 30*(2), 151–178. https://doi.org/10.1257/jep.30.2.151

Cohen, L., Malloy, C., & Nguyen, Q. (2020). Lazy prices. *Journal of Finance, 75*(3), 1371–1415. https://doi.org/10.1111/jofi.12885

Dyck, A., Morse, A., & Zingales, L. (2010). Who blows the whistle on corporate fraud? *Journal of Finance, 65*(6), 2213–2253. https://doi.org/10.1111/j.1540-6261.2010.01614.x

Ernst, H. (2003). Patent information for strategic technology management. *World Patent Information, 25*(3), 233–242. https://doi.org/10.1016/S0172-2190(03)00077-2

Fang, L., & Peress, J. (2009). Media coverage and the cross-section of stock returns. *Journal of Finance, 64*(5), 2023–2052. https://doi.org/10.1111/j.1540-6261.2009.01493.x

Fleisher, C. S., & Bensoussan, B. E. (2015). *Business and competitive analysis: Effective application of new and classic methods* (2nd ed.). Pearson Education.

Gentzkow, M., Kelly, B., & Taddy, M. (2019). Text as data. *Journal of Economic Literature, 57*(3), 535–574. https://doi.org/10.1257/jel.20181020

Gilad, B. (2004). *Early warning: Using competitive intelligence to anticipate market shifts, control risk, and create powerful strategies*. AMACOM.

Griliches, Z. (1990). Patent statistics as economic indicators: A survey. *Journal of Economic Literature, 28*(4), 1661–1707.

Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis, 21*(3), 267–297. https://doi.org/10.1093/pan/mps028

Healy, P. M., & Palepu, K. G. (2001). Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. *Journal of Accounting and Economics*, *31*(1–3), 405–440. https://doi.org/10.1016/S0165-4101(01)00018-0

Hoberg, G., & Phillips, G. (2016). Text-based network industries and endogenous product differentiation. *Journal of Political Economy*, *124*(5), 1423–1465. https://doi.org/10.1086/688176

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

Kahaner, L. (1996). *Competitive intelligence: How to gather, analyze, and use information to move your business to the top*. Simon & Schuster.

Kovach, B., & Rosenstiel, T. (2014). *The elements of journalism: What newspeople should know and the public should expect* (3rd ed.). Three Rivers Press.

Li, F. (2010). The information content of forward-looking statements in corporate filings—A naïve Bayesian machine learning approach. *Journal of Accounting Research*, *48*(5), 1049–1102. https://doi.org/10.1111/j.1475-679X.2010.00382.x

Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *Journal of Finance, 66*(1), 35–65. https://doi.org/10.1111/j.1540-6261.2010.01625.x

Loughran, T., & McDonald, B. (2016). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research*, *54*(4), 1187–1230. https://doi.org/10.1111/1475-679X.12123

Miller, G. S. (2006). The press as a watchdog for accounting fraud. *Journal of Accounting Research*, *44*(5), 1001–1033. https://doi.org/10.1111/j.1475-679X.2006.00224.x

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

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

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *Journal of Finance*, *62*(3), 1139–1168. https://doi.org/10.1111/j.1540-6261.2007.01232.x

Zhu, C. (2019). Big data as a governance mechanism. *Review of Financial Studies*, *32*(5), 2021–2061. https://doi.org/10.1093/rfs/hhy081