CONCENTRIC

DATA RISK REPORT

1H 2022

TABLE OF CONTENTS

EXECUTIVE SUMMARY

This is the 1H 2022 edition of Concentric AI's Data Risk Report. Based on information captured by the Semantic Intelligence™ solution in production deployments, we reveal how organizations create, use, and manage data.Content awareness, the subject of this report, is an increasingly important element of a defense-in-depth data protection plan. With it, organizations can implement effective least-privileges access governance and analyze the impact and implications of an attack. Learn more about the role of content awareness for security professionals.

Concentric autonomously assigns files to one of over 250 categories. Over 175 of those categories are business-critical.

TRENDS

Over 80% of an organization's data is unstructured, meaning it's embedded in the millions of financial reports, corporate strategy documents, source code files, and contracts created by CFOs, general managers, engineers, and lawyers every year. But to an IT security professional, unstructured data is still largely unseen, unexplored and insecure. Using advanced AI capabilities, Concentric processed over 550 million unstructured data files in 1H 2022, from companies in the technology, financial, oil and gas and healthcare sectors, to create this report. Semantic Intelligence categorizes documents, evaluates business criticality, and accurately assesses risk. Accuracy is critical – it's the difference between effective protection and alert fatigue caused by thousands of false positives.

Unstructured Data Trendlines, 2H 2021 to 1H 2022

AVERAGE FILES PER EMPLOYEE	7,764	Up 19%
AVERAGE BUSINESS-CRITICAL FILES PER EMPLOYEE	2,206	Up 24%
AVERAGE AT-RISK FILES PER EMPLOYEE	310	Up 24%

NEW THIS QUARTER

Education, healthcare and infrastructure continue to suffer at the hands of cybercriminals, and overshared unstructured data remain an unaddressed threat surface for many of these organizations.

- 13% of an organization's business-critical data is overshared
- On average, organizations have 598k files at-risk due to oversharing (310 files per employee) up from 498K in 2H 2021 (251 files per employee)

RESULTS IN AGGREGATE

- 550M files analyzed
- Nearly 30% of unstructured data is business-critical - that's 3.1M files in an average organization
- Over 15% of all business-critical files can be seen by internal or external users who should not have access

NEW ANALYSIS CAPABILITIES

Concentric's User360 capabilities add an additional analysis lens that can spot insider threats and identify conditions that could lead to unauthorized data exfiltration

USERS STILL MAKE CRITICAL SECURITY DECISIONS

Sharing decisions, access permission configurations and file classifications determine whether unstructured data is secure or at risk. Least-privileges data governance programs offer superior protection against ransomware and other security risks, but the IT professionals responsible for these programs lack the nuanced content knowledge needed to understand which data is critical and who should have access. End users, who better understand these issues, often fail to take data security into account when sharing conent. Attempts at more robust unstructured data security have, at least until now, required signficant investments of time, expertise and capital.

Despite these difficulties, effective data access governance and zero-trust security models that don't depend on end user decisions are critical in the fight against cybercrime. Risk Distance™ analysis from Concentric utilizes peer file analysis, enabled by advanced artificial intelligence techniques, to determine a file's business criticality and the likliehood of risk due to oversharing – without relying on end user input or complex, centrally maintained rules and policies. The remainder of this report details what we've learned.

> ”
>
> Hackers target American citizens directly every day and impact their lives at a time when we have experienced unprecedented hardships.
>
> Alejandro Mayorkas
> Secretary of the Department of Homeland Security

TAKING SHAPE

Cyberattacks target every type of organization, including schools, hospitals and infrastructure. These are the primary categories of unstructured data and what's at stake.

PRODUCT

Bills of materials, source code, design documents, and test plans

At stake: intellectual property loss, product liability, customer anxiety, strategic disclosure

FINANCIAL

Bookings, income, forecasts, pricing, invoices, trading, and tax filings

At stake: insider trading violations, compliance, loss of competitive advantage

LEGAL

Non-disclosure agreements, contracts, purchase agreements

At stake: civil lawsuits, loss of favorable supplier terms, other legal liabilities

HUMAN RESOURCES

Offer letters, stock agreements, consulting contracts

At stake: employee satisfaction, private information, higher costs

SALES

Requests for proposals, quotes, customer strategies

At stake: lost business, strategic disclosure, sales team dissatisfaction

PARTNER

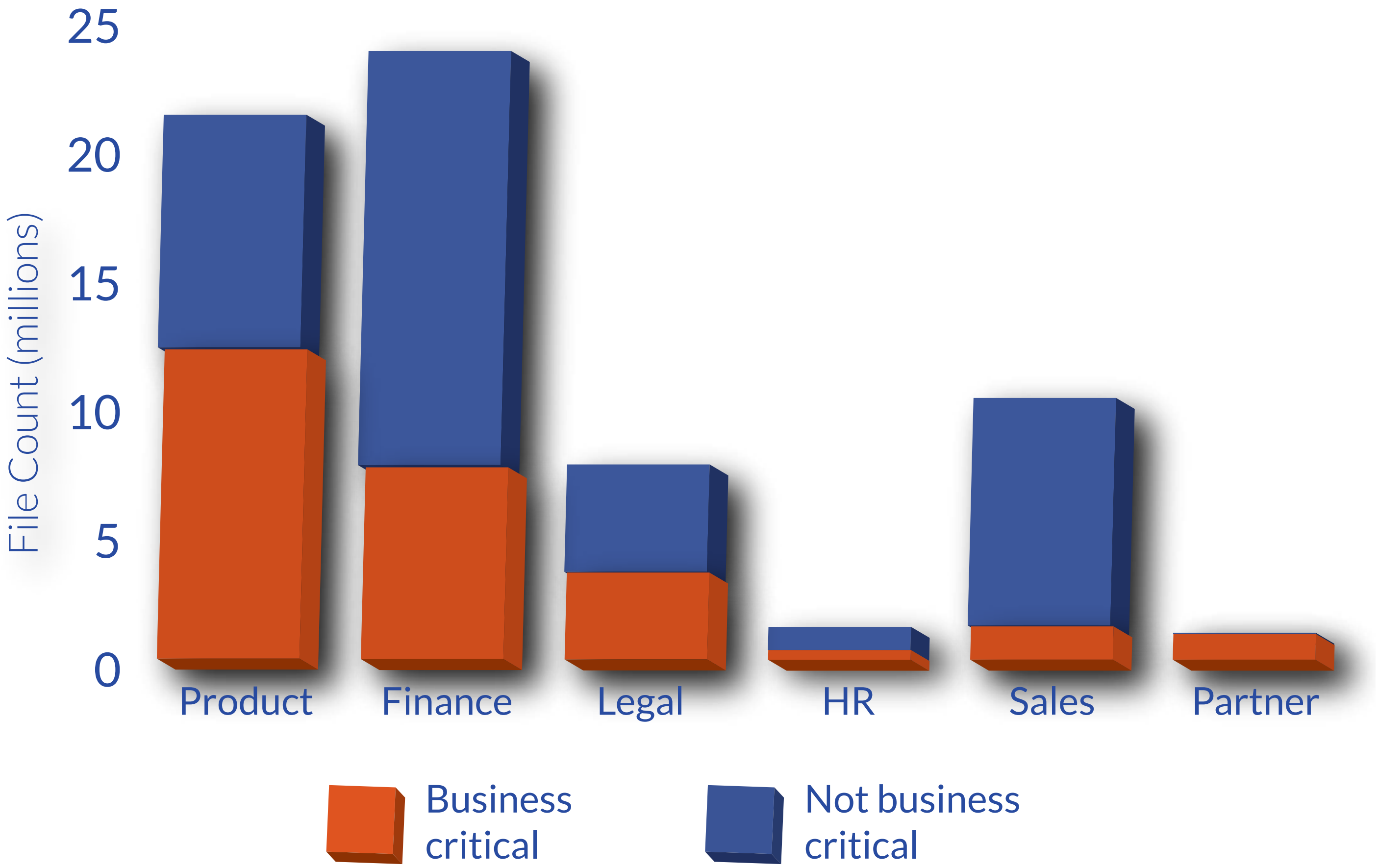
Mergers and acquisition docs, partner agreements

At stake: damage partner relationships, acquisition initiatives, insider trading

ASSESSING BUSINESS CRITICALITY

Concentric's Risk Distance™ analysis evaluates business criticality based on contextualized content, file ownership, document meta-data, presence of personally identifiable information, and peer file comparisons. Business criticality is, of course, vital to security assessment. These are the files that must not be overshared.

Files in the product and financial categories accounted for the leading share of business-critical documents (41%).



EVALUATING RISK

> END USERS MAKE SHARING DECISIONS. WHAT HAPPENS WHEN THEY GET IT WRONG?

Assessing risk – even with a fully categorized set of files - is a deceptively complex task. Appropriate sharing depends on the meaning and function of the document itself.

Concentric's Risk Distance™ analysis assesses risk by comparing each document's security characteristics to those of its peers. We evaluate:

SHARING WITH EXTERNAL USERS

Are similar documents shared with external users? Are they the same external users?

SHARING WITH GROUPS

Do peer files allow similar group access?

INTERNAL SHARING

Is internal user sharing consistent? This is tough to spot without peer file analysis – and it's critical for security.

PERSONAL EMAIL SHARING

Data shared with personal emails often indicates an insider threat.

ANONYMOUS LINK SHARING

Convenient links sharing often results in files sharing that's not appropriate or no longer needed.

MISCLASSIFIED CONFIDENTIAL FILES

Has this document been properly classified? Tags are often used by other security solutions - like DLP - to enforce policy.

MISCLASSIFIED FILES WITH PII

Is this document marked to indicate it contains PII? Classifications for PII can also help a DLP solution fence in PII to maintain compliance.

WRONG LOCATION

Are similar documents stored in a certain location? Sensitive documents in folders accessible by all employees is a common security issue.

Risk Scenarios in Business-Critical Files

EXTERNAL USER SHARING MISMATCH

Business-critical documents shared inappropriately with external users

GROUP SHARING MISMATCH

Sensitive data shared erroneously with groups

INTERNAL USER SHARING MISMATCH

Sensitive data shared erroneously with internal users

MISCLASSIFIED DOCUMENT

Misclassified documents that are accessible by the wrong personnel

UNCLASSIFIED CONTAINING PII

PII data in documents that are not properly marked/classified

WRONG LOCATION

Documents stored in locations that grant access to the wrong users

PERSONAL EMAIL SHARING

Sharing of documents with personal email accounts

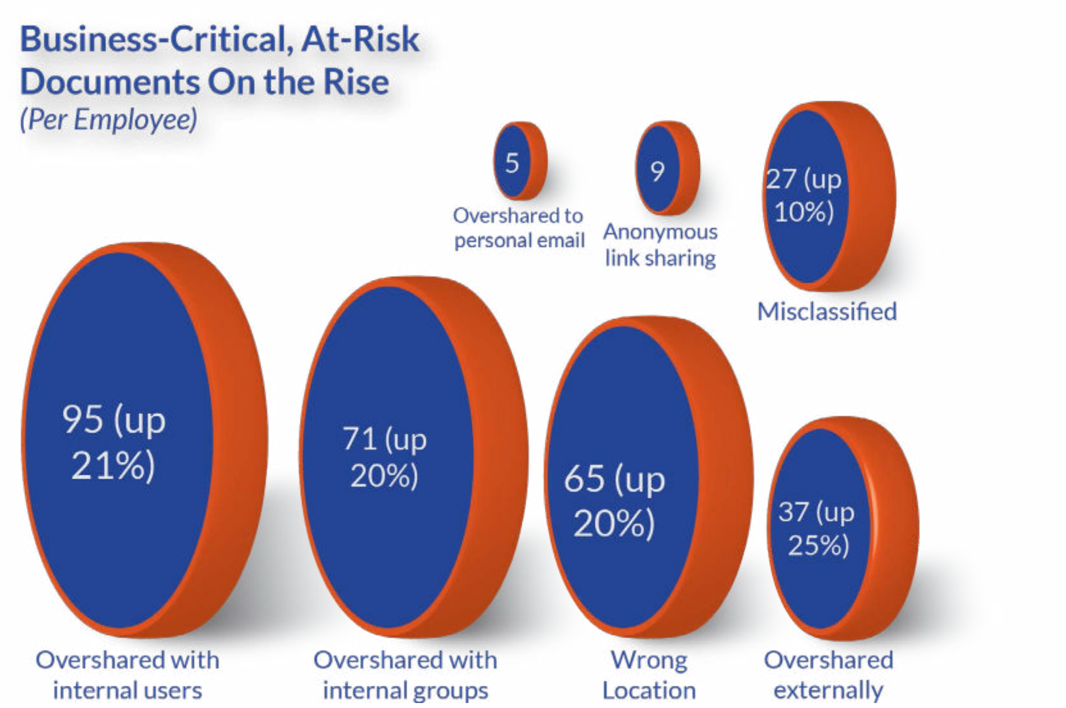
RISKY LINK SHARING

Use of link sharing to grant access that is not appropriate or no longer needed

GOING DEEP

We discovered some surprising results:

- 13% of an organization’s business critical data is overshared
- On average, each organization had 598k files at-risk due to oversharing (310 files per employee) up from 498K in 2H 2021 (251 files per employee)
- 85% of the at-risk files were overshared with users or groups within the company (flat quarter over quarter)
- 15% of business-critical files were overshared with external 3rd parties
- 280K business-critical files were erroneously classified and inappropriately accessible by other employees



DANGEROUS PATTERNS

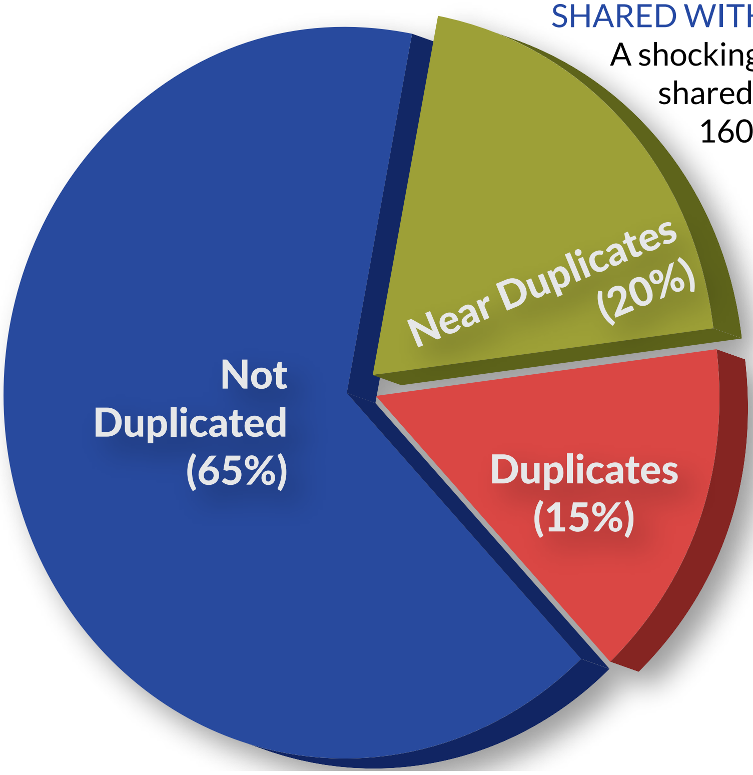
After reviewing the data, we noted some patterns that reoccur across companies, regardless of sector or company size:

NEAR-DUPLICATE FILES

Over 1 in 3 files we processed were identical or nearly identical. Near-duplicate files create multiple variant copies of sensitive information, often with insecure file permissions, prohibited locations, or improper file classifications.

SHARED WITH EVERYONE

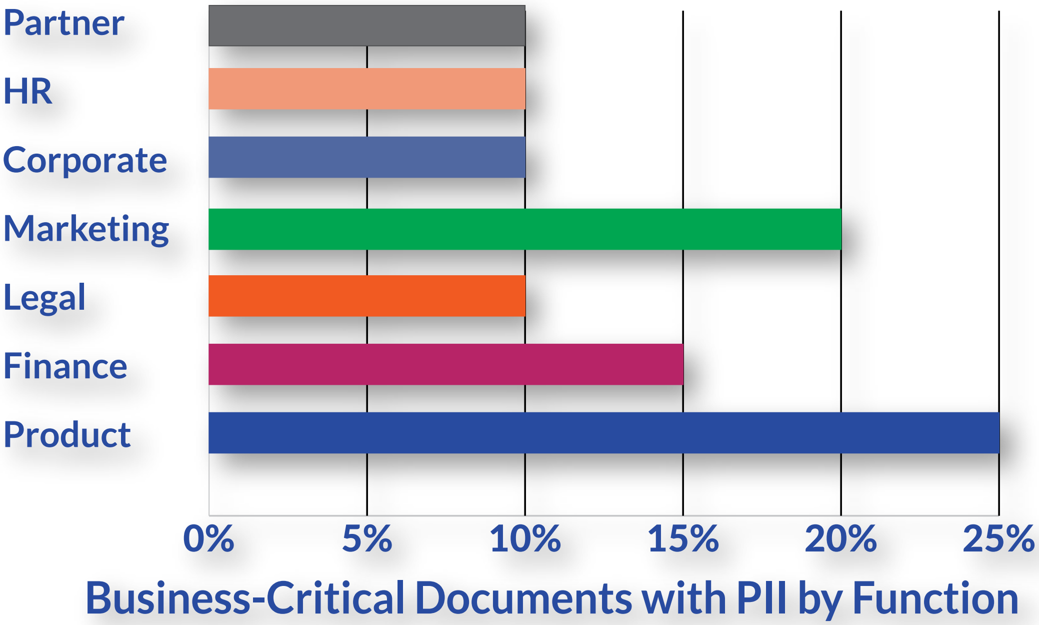
A shocking number of business-critical files were shared with everyone in the company. We found 160,078 such files.



DANGEROUS PATTERNS

PERSONALLY IDENTIFIABLE INFORMATION (PII)

PII is of interest due to privacy concerns and regulatory requirements. Increasingly, security teams use document metadata to flag PII and control document sharing and transfer. Nearly 25% of all documents containing unstructured data contained PII and were not marked appropriately.



BEWARE ORPHANED ACCOUNTS (AND OTHER SHARING MISSTEPS)

We've provided no shortage of statistics in this report. Statistics, sometimes, don't convey what's really happening on the ground. To help keep it real we offer a few specific incidents that show just how easily oversharing happens.

ORPHANED ACCOUNTS

At a professional services firm, we found 12 orphaned accounts belonging to ex-employees. Those accounts had access to 20% of the business-critical files managed by the firm.

"COMMON" FOLDER

A mid-sized financial services company uses a common folder to share non-critical documents with all employees. Out of the 100,000 files in that folder, 1,000 contained proprietary trading information or contracts. Similar documents located elsewhere in the company were highly restricted.

ENGINEERING COLLABORATION

At one high-tech firm, source code and design documents were routinely shared with everyone in the company. We found over 12,000 overshared files.

INTERDEPARTMENTAL OVERSHARING

At another financial services company, our research uncovered a sales contract shared with the Engineering and HR groups. Mis-configured access settings are common but can be as hard to find as the proverbial needle in the haystack.

CONCLUSION

Protecting unstructured data is unfamiliar territory for many security pros. Finding the intersection between business criticality, risk and urgency is no small feat, and it's the crux of the data security problem: out of the 11 million files an average enterprise has, how can we control file access without overwhelming IT teams with endless rule and policy development?

This report shows what's possible. Concentric's Semantic Intelligence™ solution uses Risk Distance™ analysis to find at-risk data and protect organizations from theft and loss. And now, with our new ransomware defense capabilities, establishing cybersecurity hygiene for data access governance is within reach.

Business Criticality
Risk
Urgency

CONCENTRIC

Eighty percent of corporate data is found in the files and documents employees create and use every day. Concentric discovers and categorizes this unstructured data to protect intellectual property, financial documents, PII/PCI content and proprietary business information (strategy plans, product roadmaps, contracts, blueprints) wherever it's stored. Our Semantic Intelligence™ solution uses deep learning to develop a semantic level understanding of a document's content to discover business sensitive data, surface risks, and remediate issues without relying on upfront rules or complex configuration.

www.concentric.ai
Twitter: @IncConcentric
LinkedIn: linkedin.com/company/concentricinc

19548 Chardonnay Ct
Saratoga
CA 95070