Mapping the Dark Data Universe

10 Places
Dark Data Resides

- **Hard drives, storage systems, & arrays**
  The average business stores 10PB of data or 2.3 billion files, and that amount is growing by 39% each year.[3]

- **Cloud services**
  Over 1 exabyte (1024 petabytes) of dark data is stored in the cloud.[4]

- **Image & video files**
  The average organization stores 3.3 billion image files and 1.4 billion video files.[5]

- **Text files**
  The average organization stores 800 million text files.[6]

- **Database files**
  The average business stores 1.1 billion database files—many in non-relational databases, such as NoSQL.[7]

- **Emails & attachments**
  Over 333 billion emails are sent and received each day.[8]

- **Business collaboration tools**
  Nearly 80% of workers are using collaboration tools for work in 2021, up from just over half of workers in 2019, according to Gartner.[9]

- **Employee endpoints**
  20% of employees have used personal devices to store business information in the last year.[10]

- **Customer support systems**
  Organizations store information on customers and prospects across disparate databases, customer support systems, server files, call logs and records, call center recordings and geolocation information.

- **IoT sensors**
  The amount of data generated by IoT devices is expected to reach 73.1 ZB (zettabytes) by 2025.[12]

- **...and many more**

>80% [1]

The average amount of data that is dark for any given organization. Meaning it's unknown, unstructured – and unprotected.

Why is there so much dark data?

Respondents to a recent survey said the obstacles to surfacing dark data include:[2]

- The sheer volume of dark data
- Lack of necessary skill sets
- Lack of resources
- Difficulty in coordinating across departments
- No one is dedicated to finding data the organization possesses
- Difficulty in coordinating with data–generating third parties
- Lack of control over data-generating devices and apps
- Lack of interest from organization leaders

44 zettabytes [11]

or, **44 trillion gigabytes**, is the estimated amount of data contained in systems worldwide.

$29 million [13]

The average annual cost of managing dark data that hasn't been modified in more than 3 years.

The Hidden Challenges of Dark Data

- **Cyber attack risks**
  The average cost of a data breach in the United States in 2021 was $4.24 million.[14]

- **Compliance risks**
  Unsecured dark data poses legal, security, and compliance risks. Organizations with high compliance failures pay an average of $2.3 million more for data breaches.[15]

- **Missed opportunities**
  Organizations that are able to leverage all of their data can expect to see productivity gains of $30 billion over their peers.[16]

- **Increased (and wasted) storage costs**
  Research from Gartner estimates data center server spend will grow from $80 billion in 2019 to $111.6 billion by 2025.[17]

$3.3T [18]
How much not knowing the answers hidden in dark data costs companies globally.

287 [19]
The average number of days it takes to identify and contain a data breach.

77%
Percentage of respondents in a survey by Splunk who agreed that finding and capturing dark data is a priority.

ROI of 2.18
IT leaders surveyed estimate that investing in better data management will result in an average ROI of 2.18 for every dollar spent.[20]
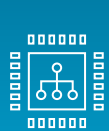
Bring Your Dark Data Universe to Light

Take these steps to explore, classify and bring your dark data universe to light.

01 Discover
Identify all data and understand its sources using an **accurate data discovery tool**.

02 Classify
Classify and determine the data's importance with **persistent data classification**.

03 Remediate
Set retention and secure deletion procedures with **intelligent data remediation**.

Don't Get Lost in the Dark...

Assess your environment to improve visibility and efficiency with a data risk assessment from Spirion. This assessment will:

- Identify compliance violations
- Find security gaps
- Uncover vulnerabilities
- Detect malicious activity

Contact a data security expert at expert@spirion.com or visit spirion.com

SPIRION IS
step one
FOR SENSITIVE DATA GOVERNANCE

Sources:
1. https://www.ibm.com/blogs/cloud-archive/2015/11/future-of-cognitive-computing/
2. "The State of Dark Data" by Splunk.
3. "Dark Data: Are You at Risk?" research from Veritas.
4. https://www.neowin.net/news/research-firm-over-1-exabyte-of-data-is-now-stored-in-the-cloud/
5. "Dark Data: Are You at Risk?" research from Veritas.
6. "Dark Data: Are You at Risk?" research from Veritas.
7. "Dark Data: Are You at Risk?" research from Veritas.
8. https://www.statista.com/statistics/456500/daily-number-of-e-mails-worldwide/
9. https://www.gartner.com/en/newsroom/press-releases/2021-08-23-gartner-survey-reveals-44-percent-rise-in-workers-use-of-collaboration-tools-since-2019
10. "Dark Data: Are You at Risk?" research from Veritas.
11. https://www2.deloitte.com/us/en/insights/focus/tech-trends/2017/dark-data-analyzing-unstructured-data.html
12. https://www.businesswire.com/news/home/20190618005012/en/The-Growth-in-Connected-IoT-Devices-is-Expected-to-Generate-79.4ZB-of-Data-in-2025-According-to-a-New-IDC-Forecast
13. https://www.ibm.com/security/data-breach
14. https://www.ibm.com/security/data-breach
15. https://www.ibm.com/security/data-breach
16. https://www2.deloitte.com/us/en/insights/focus/tech-trends/2017/dark-data-analyzing-unstructured-data.html
17. https://www.gartner.com/en/documents/4008896
18. Marketplace Research: Data Integration and Exploration
19. https://www.ibm.com/security/data-breach
20. The Value of Data Study, conducted by Vanson Bourne for Veritas.