
Dark data under the spotlight: How organizations can turn liability into asset

By David Weldon

October 21, 2019, 3:30 a.m. EDT



Advances in the sophistication of unstructured data analytics and text analytics means that dark data increasingly has the potential to be a business asset, but the real challenge lies in determining its business value.

To get a better understanding of dark data and how it can rise from liability to asset, Information Management recently spoke with Kon Leong, chief executive officer at ZL Technologies, Inc. According to Leong, bringing dark data to light holds the potential for eliminating repetitive human effort and increasing the productivity of the entire organization.

Information Management: What exactly is dark data, where can it be found, and what should data management professionals know about it?

Kon Leong: By definition, dark data is the unmanaged enterprise content lurking in the shadows. Content that goes unmanaged, likewise, is difficult to monitor and the inability to monitor usually amounts to the inability to notice when information has been replicated, leaked, tampered with, lost or stolen. These potentially fatal outcomes for enterprises reveal why data management professionals must find ways to understand, and ultimately manage, dark data.

IM: What are the security risks associated with dark data?

Leong: In looking at our biggest security risks, it is not as much what is outside the firewall – which we valiantly protect and invest in securing – but it's really the risk of insider threats that can cost companies. It's the data that is sitting unsecured within the organization that poses the number one risk.

This isn't to say that workers are intentionally perpetuating risk: dark data is simply a byproduct of the chaotic human work process. The majority of security breaches perpetuated by employees are not malicious in nature; accidental disclosure of sensitive content is far more common than corporate espionage.



Alex Tihonov/Alex Tihonov - Fotolia

So the biggest risks are due to accidental copy and transfer of sensitive information, as well as improper access control to documents and files. As long as there is dark data in the work environment, these security risks will largely be unnoticed until something catastrophic occurs, which might mean it is too late.

IM: Are there ways that organizations can reduce those risks?

Leong: Cleaning up dark data means shining light into the darkest corners of enterprise information. Often, these are found within the organization's file share environment where workers create, modify, store and collaborate on documents and other files.

Due to the sheer volume of data created in these file shares, most organizations don't even know what they have. And when they do try to understand what's inside, they don't know where to start. Often, the extent of knowledge regarding the file share is limited to the total amount of storage being used, and the number of users.

The reason it presents a risk is twofold. One, because of its scale: more data and more users means more cumulative risk. Two, because of the rapidly-changing and messy nature of human collaboration: often, it's difficult to tell the difference between a final "important" document and the incomplete versions that preceded it. The end result is a living ecosystem that is always changing.

Reducing risk means having a systematic way to eliminate duplicate copies, identify sensitive information, control the lifecycle of content and monitor use. Due to the scale of modern data, a manual approach to management is not feasible; there needs to be an automation in order to scale.

So modern efforts at information governance generally attempt to pool information into a single "lake" in which items can then be managed based on author, document age or any other number of variables. The problem is that most organizations are struggling to implement this management aspect, and many tools on the market are not sophisticated enough to control items with the granularity that businesses need.

IM: How could dark data potentially be of value to an organization instead?

Leong: To date, dark data has often been viewed – rather fairly – through the lens of risk reduction. Nobody wants a skeleton in the closet when it comes time to litigate or answer the demands of regulators.

However, the current advances in the sophistication of unstructured data analytics and text analytics means that dark data increasingly has potential as a business asset. The challenge therein, however, is extracting that value. To do so requires ongoing management of data.

Of course, saying that "dark data" itself has value is a paradox. Generally, dark data has little immediate value. However, dark data has immense potential to create value. Once it has been managed, de-duplicated, and cleansed it represents a corpus of human knowledge within the organization.

Dark data, in essence, is the cluttered human mind of the business. The challenge is curating that knowledge into useful resources. Once the information has been managed and analyzed, it offers immense insight into human work patterns, subject matter expertise, communication networks, influencers and business processes. Bringing dark data to light holds the potential for eliminating duplicative human effort and increasing the productivity of the entire organization.

Moving forward, there will likely be a shift in perception regarding the potential of dark data. Long gone are the days when organizations systematically discarded content that wasn't proven to be a formal business "record." Rather, the big data era tends towards hoarding – for better or worse.

With the increase in sophistication of both analytics capabilities and governance capabilities, we will likely see a renaissance for this data that is currently being viewed as liability.

David Weldon Editor-in-chief, Information Management



 **REPRINT**

For reprint and licensing requests for this article, [click here](#).

[Data management](#)[Analytics](#)[Data visualization](#)

TRENDING

FINTECH

Why Scotiabank is building 'digital factories'

The Toronto-based bank has hired 1,200 people across five countries in North and South America to focus on digital features.

By [Will Hernandez](#) [October 22](#)

DATA MANAGEMENT

Firms that excel at scaling tech innovation are rewarded with greater profits

Organizations that excel at scaling technology innovation generate double the revenue growth, according to a new report from IT services and consulting firm Accenture.

By [Bob Violino](#) [October 22](#)

DATA SECURITY

7 views on the important lessons of Cybersecurity Awareness Month

This month is National Cybersecurity Awareness Month, a time for organizations and individuals alike to think about their data, likely security threats and best practice defense strategies.

By [David Weldon](#) [October 18](#)

MORE FROM INFORMATION MANAGEMENT

ANALYTICS

Poor data access putting many analytics projects at risk

By Bob Violino October 18

ARTIFICIAL INTELLIGENCE

How BNY Mellon is going further on AI

By Penny Crosman October 17

ARTIFICIAL INTELLIGENCE

AI and machine learning growing as an industry force, experts say

By Joseph Goedert October 16

CLAIMS

Auto claims leans into AI, automation

By Pat Speer October 16

DATA STORAGE

Market for data warehouses expected to double over next five years

By Bob Violino October 16

EHR DOCUMENTATION

Practices need to vastly improve data accuracy in records, attorney says

By Joseph Goedert October 15

Like what you see? Make sure you're getting it all

Authoritative analysis and perspective for data management professionals.

[SIGN UP](#)



[About Us](#)

[Contact Us](#)

[Site Map](#)

[Subscription Agreement](#)

[Content Licensing/Reprints](#)

[Advertising/Marketing Services](#)

[Information Management's Glossary](#)

[RSS Feed](#)

[Privacy Policy](#)

© 2019 SourceMedia. All rights reserved.