

1,397 views | Oct 27, 2019, 10:14am

# What You Need To Know About Dark Data



**Tom Taulli** Contributor 

Entrepreneurs

*I write about tech & finance.*



multiple layers binary code background. number one and zero in pattern moved upward. GETTY

The concept of **dark data** sounds ominous, even sinister. But it is very important in the technology world. “To make it more relatable, dark data is like all of the photos on your devices,” said Sky Cassidy, who is the CEO of [MountainTop Data](#). “Most of them will never be used or even viewed again, but they are there. So as for dark data, it’s all the information companies collect in their regular business processes, don’t use, have no

plans to use, but will never throw out. It's web logs, visitor tracking data, surveillance footage, email correspondences from past employees, and so much more."

For most companies, there is usually an enormous amount of dark data. According to Rahul Telang, who is the professor of information systems at [Carnegie Mellon University's Heinz College](#), it's about 90% or so.

While dark data may never be used or be useful for many organizations, it's something that should not be ignored. "One key area of challenge is to manage dark data in order to minimize risks and liabilities in information governance, that is, regulatory compliance, litigation, records-keeping, privacy and records management," said Kon Leong, who is the CEO of [ZL Technologies](#). "Separately, while analytics is a very active area of focus for data managers, it is now critical that analytics and governance be addressed simultaneously going forward. New privacy laws, for example, now cover all data, including data in analytics repositories."

Today In: [Small Business](#)



But dark data should not just be about handling regulatory issues. This type of information may ultimately prove quite useful for deriving insights for managers. We are already seeing this with software that helps automate and streamline operations, such as with RPA (Robotic Process Automation).

"Cognitive automation is the answer," said Prince Kohli, who is the CTO of [Automation Anywhere](#). "By adding structure to unstructured content, cognitive RPA helps you automate invoice, purchase-order, and mortgage application processing—all of which rely on the dark data stored in documents, images, emails, and more. At Automation Anywhere, we believe that within five years, knowledge workers will be freed from the task of extracting information from unstructured content. They will then be empowered to do what they do best: make decisions, handle exceptions, and interact with customers, partners, and each other to advance business objectives."

Then what are some of the best practices with dark data? What can be done to get the most value from it?

Well, a first step is to use data classification so as to get a sense of what you are working with. “We are seeing a number of vendors from the backup and recovery market segment come to market with solutions to help better do this and make it easier for re-use,” said Christophe Bertrand, who is a Senior Analyst at [Enterprise Strategy Group \(ESG\)](#). “Taking a holistic perspective is necessary, and leveraging existing processes that already move and manage data is critical.”

After this, you can look at what data has some potential and what is essentially unnecessary. “You can reorganize the data logically in a repository so employees can find the documents faster,” said Ilia Sotnikov, who is the VP of Product Management at [Netwrix](#). It’s also a good idea to put together a data strategy as well as a governance policy.

“Working with dark data is not a one-time project,” said Sotnikov. “Over time, you need to improve it.”

*Tom (@ttaulli) is the author of the book, [Artificial Intelligence Basics: A Non-Technical Introduction](#).*

*Follow me on [Twitter](#) or [LinkedIn](#). Check out my [website](#) or some of my other work [here](#).*



**Tom Taulli**

Tom (@ttaulli) is the author of the upcoming book, [Artificial Intelligence Basics: A Non-Technical Introduction](#).