

Dark Data

The information assets that organizations generate, collect, process & store during regular business activities, but generally fail to use for other purposes like analytics, monetizing etc. This unanalyzed data which is a subset of big data contains undiscovered, important insights and represents an opportunity lost. If not handled well, continuous storage of this data puts the organization at risk, especially when the stored data is sensitive.



Categories

- ◆ Customer Information
- ◆ Log Files
- ◆ Account Information
- ◆ Notes Or Presentations
- ◆ Email Correspondences
- ◆ Financial Statements
- ◆ Raw Survey Data
- ◆ Analytics Reports
- ◆ Previous Employee Information
- ◆ Old Versions Of Relevant Documents

Reasons

- ◆ Shortage Of Tools To Capture And Unlock Dark Data
- ◆ Too Much Data But Not Enough Analytics
- ◆ Able To Access Structured Data Only
- ◆ Missing Or Incomplete Data
- ◆ Poor Quality Data
- ◆ Lack Of Business Intelligence Tools To Clean Or Organize Data

Trends

- ◆ According to a IBM report, 80% of all the data is dark and unstructured
- ◆ A report by Seagate and IDC found 68% of the data available to enterprise goes unleveraged
- ◆ By 2025 the amount of data stored around the world will grow to 175ZB from 33ZB in 2018, i.e., we will be managing over 4 times the current volume of dark data in just 5 years
- ◆ By 2025 nearly 80% of the world's data will be stored in the core and edge, up from 35% in 2015