Taking A Big Byte Out of Dark Data

Vickey Seehusen

Edgar Maldonado

Businesses and individuals continue to create massive amounts of data, and this exploding data volume is often overlooked by most business leaders. Veritas, an enterprise data protection company, predicts that by 2020, organizations around the world are on course to waste more than 3.3 trillion dollars storing data that is considered redundant, obsolete and trivial (ROT data) (Smith, 2016). Most of the storage data has not been carefully indexed, so its value has not yet been identified, the term for this kind of information is Dark Data This paper classifies the concerns about maintaining dark data according to current literature. The paper also presents a taxonomy for classifying dark data; some dark data might be useful for business decision making and predictive analysis and the non-useful data could then be destroyed, freeing up corporate computing and storage assets. Managing and indexing dark data is an understudied topic in the academic world and thus solutions to the problems of dark data would benefit from additional research.