

Ensuring the protection of sensitive information is of vital importance to both the business and government sectors. Many approaches to this problem have focused on the use of text classification and information retrieval techniques to identify text with sensitive information. While these techniques can work, and are fast, they lack the ability to provide users with the rationale as to why a piece of text has been determined to be sensitive. This paper outlines an approach that leverages domain-specific information extraction to extract facts from text and map those to an existing domain ontology so to imbue the extracted facts with further context. It then uses logical inferencing to reason on the sensitivity of the text. This provides users the exact parts of text that are sensitive, the rationale for why something is sensitive, and the provenance that can point to where a sensitivity guideline came from.