Investment Bank Case Study: Records and Retention Management

This major investment bank was found in contempt of court and massively fined for their incoherent approach to records and retention management. Up to this point the prevailing approach had been to allow data stewards to tag documents and systems with record classification information to aid in the process of discovery and disposition.

Our subsequent analysis revealed what many suspected. Of the almost 10,000 application databases and over 60,000 content repositories, approximately one-half of one percent had been classified.

Our sponsor had the intuition that context was the key to improving this problem. There was a wealth of contextual information, but it was in various different systems, uncoordinated and in many cases shrouded in acronyms and arcane terms.

We extracted a great deal of contextual information. It turns out knowing who set up a repository, what department they work in, what cost center they charge it to, how they named it, and where they put it are valuable clues as to what the repository contains. But these clues can only make sense with a bit more mining.

To begin, we harvested their financial reporting structure, the cost center structure, and all the employees. For each, we also got as much narrative as we could. We got the division and department description and mission, the reason for setting up the cost center, and the job description for each employee. We unpacked the acronyms. We loaded all this into a knowledge graph.

With some very lightweight NLP we were able to get accurate classification for about 25% of the repositories based on this information. Quite a difference from the one-half of one percent. This was enough to launch a major effort that now uses machine learning to mine deep learning that allows knowledgeable analysts to classify with even higher degrees of accuracy and completeness.

Contact Us:

Overcome integration debt with proven semantic solutions.
Contact Semantic Arts, the experts in data-centric transformation, today!
CONTACT US HERE

Address: *Semantic Arts, Inc.*
123 N College Avenue Suite 218
Fort Collins, CO 80524

Email: *info@semanticarts.com*

Phone: *(970) 490-2224*