Case Study

Tackling Redactions: Automating PII Redactions using Artificial Intelligence

Situation

TCDI's client, a major corporation, faced a court order to publish almost 4,000,000 documents to a public repository. Part of this mandate required a thorough review of those documents for various reasons, including the redaction of Personally Identifiable Information (PII). Given the immense volume and looming deadline, TCDI leveraged their strategic partnership with TackleAI, a leader in autonomous redactions, to tackle the PII redaction process efficiently.

Resolution

Understanding that not all documents would require redaction, TCDI's team initiated their process by using in-house PII identification software to scan the entire dataset and flag documents likely needing redaction. After quality control (QC) review, TCDI identified over 450,000 documents that potentially contained PII, indicating the need for redaction.

TCDI then collaborated with TackleAI, sharing the identified documents and the PII redaction protocol that was created. Maintaining open communication was crucial, as it allowed TCDI's team to address any questions regarding PII and to receive timely updates on progress.

TackleAI used its proprietary Advanced Content Extraction (ACE) process to read the documents. Blending elements of OCR, Computer Vision, and other enhancements. ACE allows TackleAI to work with both native and static images, even of degraded quality, and identify embedded elements missed by traditional approaches.

TackleAI then calibrated its models to adhere to the redaction protocol and ran against the data set to identify PII within. Using a combination of machine learning, computer vision and quantum mesh probability, TackleAI's proprietary AI was able to automate the entire identification and redaction process.

Once TackleAI completed the redactions, the redacted images were reloaded into TCDI's database for further review. TCDI's review team conducted statistical sampling of the redacted documents at a 95% confidence level with a 5% margin of error. The results were impressive— no missed redactions and only a single instance of over-redaction in that sample set, which was quickly addressed with TackleAI to ensure it was an isolated case.

Additionally, TCDI sampled the documents that TackleAI did not flag for redaction (the null set), identifying some instances where personal email addresses were missed due to degraded image quality in the production set provided to TackleAI. By using the extracted text available in the database, TCDI pinpointed the documents requiring additional redactions and worked with TackleAI to update them accordingly.

Impact

Traditionally, PII redaction requires large teams of reviewers to manually review and redact documents, a process that is both time-consuming and costly. By utilizing TCDI's process-driven AI solution in partnership with TackleAI, they redacted PII in 439,000 documents within just five weeks— compressing the timeline by two months and delivering a 38% cost saving for their client.

This collaboration exemplifies the power of AI in transforming document review processes, ensuring accuracy, efficiency, and significant cost savings for TCDI's clients.

We value our clients and are proud of the partnerships we have nurtured over the years. We honor their need for discretion. Client references are available upon request with a fully executed NDA.