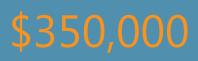
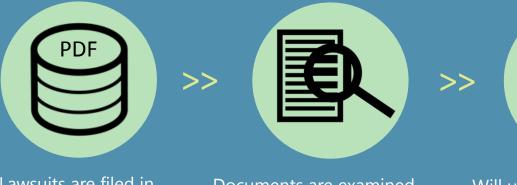
Federal Litigation Database Data Extraction and Manipulation

Information Opportunity: Proactively mitigate legal risk



cost of an average lawsuit (losses, expenses, attorneys' fees)



Lawsuits are filed in federal database

Documents are examined for relevant keywords Will use findings to create subject matter word clusters

## Methodology

## Use Python

- Manipulate files
- Extract case facts
- Sift contents





## Next Steps

- Extracted content
  will be handed off
  for analysis
- Data to be used for a system to prevent litigation



Karen Mason, MSIM Acknowledgements: Nick Brestoff



Creative Commons Attributions for Icons: Database – Anton Outkine, Search – Gianni Dolce Merda, Structure – Nicolay Necheuhin, Filter – TDL London, Relay – Andrew McKinle