A health care claim processing application is introduced which processed both structured and unstructured information associated with medical insurance claims. The application makes use of a natural language processing (NLP) engine, together with application-specific knowledge, written in a concept specification language. Using the NLP techniques, entities and relationships that act as indicators of recoverable claims are mined from management notes, call centre logs and patient records to identify medical claims that require further investigation.

Text mining techniques can then be applied to find dependencies between different entities, and to combine indicators to provide scores to individual claims. Claims are scored to determine whether they involve potential fraud or abuse, or to determine whether claims should be paid by or in conjunction with other insurers or organizations. Issues related to the design and operation of the application are discussed, specifically the use of rule-based techniques which provide a capability for deeper analysis than traditionally found in statistical text-mining techniques.