Web crawler crawls list of well known news sources (initially non-local) and the text is stored on our server for analysis.  
  
Each article has a title, URL, source (news outlet), date, category (political, financial, etc.), subcategories (editorial, etc.), text, and heuristics (including interest level).  
  
Crawler periodically returns to the source and diffs the content against the stored text for the article to check whether the article has been updated.  
  
Stream (probably a different term?): User can search a story by keywords. The user can also provide a URL or title to an article if they want to find more news about the same story or the timeline for the story, etc.