

Academic fulltext search engine Egothor has recently become starting point of several thesis aimed on searching. Until now, there was no solution available to provide robust set of web content processing tools. This master thesis is aiming on design and implementation of distributed search system working primary with internet sources. We analyze first generation components for processing of web content and summarize their primary features. We use those features to propose architecture of distributed web search engine. We aim mainly to phases of data fetching, processing and indexing. We also describe final implementation of such system and propose few ideas for future extensions.