**Abstract**

The purpose of this thesis is to conceive and implement a distributed web crawler information in cooperation with the company Immowelt AG. This web crawler is applied in the crawling stage to retrieve real estate (planning permissions on the local authorities website in PDF-Format).

This web crawler should perform better in terms of crawling quantity and quality comparing to traditional single-thread crawler. This can be realized with the use of various technologies on cloud platform like Amazon Web Services (AWS). Furthermore the use of container technologies should be utilized to increase the flexibility of continuous deploy and scaling of the web crawler.

The result of this thesis should be a prototype of a web crawler that delivery