atitit.基于内容标记的内容结构化提取

我们从网爬虫抓取内容的时候儿...title 与 内容的提取是个问题...

如何识别title与内容??

使用标记技术...

Title 开头atitit

Content startwith tit.... Endwith --atiend --attilaxend