TITLE

# INTRODUCTION

Blablabla

## Runtime Environment

Operation System: Windows 2000 or later (For Folder Create)

Framework: .NET Framework 2.0

Requested Execution Level: Require Administrator

Network Connection: Require

## Compile Environment

IDE: Microsoft Visual Studio 2010 or later

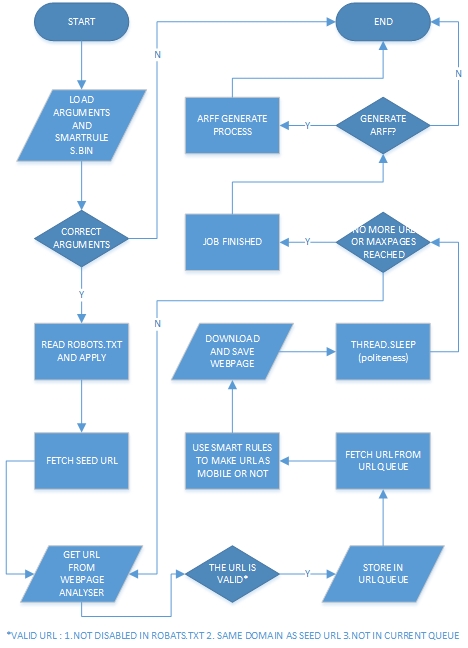
Compile Without IDE: csc.exe

# IMPLEMENTATION

## START APPLICATION

Crawl [-politeness <seconds>] [-maxpages <pages>] seed\_url

## FLOW CHART



## DATA STRUCTURES

clsURL {string URL, string FileSave, bool isMobile}

clsQueue {Queue UrlQueue, int GetQueueNumber}

clsSmartRule {string Website, string PathBelongToMobile}

## CLASSIFICATION

When Url contains certain strings which define in ‘SmartRules.bin’, the Url can be marked as ‘Mobile’.

‘SmartRules.bin’ is an ‘.ini’ configuration files which can be read by application and edited by users.

## CLASSIFICATION STRUCTURE

[http://www.theage.com.au]

MOBILEURL=MOBILE.IPHONE.IPAD

[WEBSITE URL]

MOBILEURL=Trigger1.Trigger2…

If a url contains trigger string, then ‘crawl’ will mark this url as Mobile. Users can use ‘Correct Classifier’ to improve the performance.

## ARFF FILE STRUCTURE

@RELATION Crawler

@ATTRIBUTE URL string

@ATTRIBUTE Length numeric

@ATTRIBUTE Lable string

@ATTRIBUTE StoreFile string

@ATTRIBUTE class {Mobile,NotMobile}

## APPLICATION PERFORMANCE

Use .NET delegate function to prevent application stop responding.

# CONCLUSION

## Future Improvement

## Acknowledgement