

# FPT UNIVERSITY

## Capstone Project Document

Anki Pan

AKP Team		
Project team	Bui Hong Nhung	SE02082
	Vu Hong Thai	SE02654
	Le Thanh Binh	SE02917
	Nguyen Duc Huy Hoang	SE02651
	Pham Trong Nghia	SE02781
	Ngo Thi Mai	SE02218
Supervisor	Mr. Tran Binh Duong	
Project code	AKP	

Ha Noi 14<sup>th</sup> Jul, 2015

## Acknowledge

*We specially thank to Mr. Tran Binh Duong, our supervisor, for his support through out this project. His instruction and motivation has been lessons for us not only about technique, but also about how to work as a team in a real project.*

*We also thank to our families and friends. Their encouragement has helped us to overcome many pressure to complete our product.*

## TABLE OF CONTENTS

I. INTRODUCTION .....	1
1. PROJECT INFORMATION .....	1
2. PROJECT TEAM .....	1
3. INITIAL IDEAL.....	2
3.1. Background .....	2
3.2. Existing Products.....	3
3.2.1. Google translate.....	3
3.2.2. CamDictionary .....	6
3.2.3. JDict .....	9
3.2. Our Proposal .....	11
II. PROJECT MANAGEMENT PLAN .....	13
1. PROJECT OVERVIEW .....	13
1.1. Project description.....	13
1.2. Scope.....	13
1.3. Standard Objectives.....	13
1.4. Milestones and Deliverables.....	14
1.4.1. Milestones .....	14
1.4.2. Deliverables .....	15
2. PROJECT ORGANIZATION.....	17
2.1 Software Process Model .....	17
2.2 Project Lifecycle .....	17
2.3 Roles and Responsibilities .....	18
3. TOOLS AND INFRASTRUCTURES .....	19
3.1. Hardware .....	19
3.2. Software .....	20
4. SCHEDULES .....	21
4.1. Detail Schedules.....	21
5. COMMUNICATION MANAGEMENT .....	25
5.1. Stakeholders and Contacts .....	25
5.2. Communication Management Approach.....	26
5.3. Communication Requirements .....	26
5.4. Communication Method and Technologies.....	26
5.5 Communication Matrix.....	27
6. RISK MANAGEMENT <sup>l</sup> .....	29
6.1. Risk Management Approach.....	29
6.2. Risk Identification.....	29
6.3. Risk Qualification and Prioritization .....	29
6.4. Risk Monitoring .....	29
6.5. Risk Register .....	30
6.5.1. Risk description .....	30
6.5.2. Probability – Impact matrix .....	31
6.5.3. Risk response plan & Risk status.....	32

7. QUALITY MANAGEMENT .....	34
7.1. Quality Management Overview.....	34
7.1.1. Organization, Responsibilities, and Interfaces .....	34
7.1.2. Tools, Environment, and Interfaces .....	34
7.2. Quality Planning.....	34
7.2.1. Define Project Quality .....	34
7.2.2. Measure Project Quality.....	35
7.3. Quality Assurance.....	36
7.3.1. Analyze Project Quality .....	36
7.3.2. Improve Project Quality.....	37
7.4. Quality Control .....	37
7.5. Action Plan.....	38
III. SOFTWARE REQUIREMENT SPECIFICATION .....	40
1. INTRODUCTION .....	40
1.1 Purpose.....	40
1.2. Scope.....	40
1.3. Overview.....	40
2. OVERALL DESCRIPTION .....	40
2.1. Scope of work.....	40
2.2. Assumptions.....	41
3. FUNCTIONAL REQUIREMENTS .....	42
3.1. Android application.....	42
3.1.1. User Case Diagram.....	42
3.1.2. Functions.....	43
3.2. Server .....	76
3.2.1 User case diagram .....	76
3.2.2. Functions.....	77
4. NON – FUNCTIONAL REQUIREMENTS .....	102
4.1. Usability.....	102
4.2. Availability.....	102
4.3. Performance .....	102
4.4. Interface .....	102
4.4.1. User interface.....	102
4.4.2. Hardware interface.....	102
4.4.3. Software interface.....	102
4.5. Security .....	102
IV. SOFTWARE DETAIL DESIGN .....	103
1. ARCHITECTURE DESIGN .....	103
1.1. Introduction.....	103
1.1.1. Purpose.....	103
1.1.2. Scope .....	103
1.1.3. Overview .....	103
1.2. Description of Architecture Design.....	104
1.3. Architecture Goals & Constraints.....	104
1.3.1. Technical Platform.....	105
1.3.2. Communication.....	105
1.3.3. Security .....	105
1.3.4. Persistence .....	105
1.3.5. Reliability/Availability.....	105

1.3.6. Performance .....	105
1.4. User Case View .....	105
1.5. Deployment View .....	107
1.6. Implementation View.....	108
1.7. Detail Design of Application .....	108
1.7.1. Overview .....	108
1.7.2. Design.....	109
1.8. Detail design of Server.....	162
1.8.1. Overview .....	162
8.1.2. Detail design .....	163
1.9. AKP image processing library .....	187
1.9.1. Purpose.....	187
1.9.2. Tesseract' shortcoming .....	188
1.9.3. Tesseract's recognition of Japanese.....	190
1.9.4. AKP Image Processing Library .....	191
1.9.5. AKP image processing library's limitation.....	201
1.10. Size And Performance.....	201
1.11. Quality.....	201
1.11.1. Scalability:.....	201
1.11.2. Reliability / Availability:.....	201
1.11.3. Portability: .....	201
1.11.4. Security: .....	201
2. DATABASE DESIGN .....	202
2.1. Introduction.....	202
2.1.1. Purpose.....	202
2.1.2. Overview .....	202
2.2. Database.....	203
2.2.1. Application database .....	203
2.2.2. Server database .....	206
2.3. Code Design.....	208
2.3.1. Japanese_dict.content.....	208
2.3.2. Kanji.content.....	208
2.3.3. User.password.....	208
3. INTERFACE DESIGN .....	209
3.1. Introduction.....	209
3.2. Interface Design of Application .....	209
3.2.1. Structure and Layout.....	209
3.2.2. Screen Description.....	211
3.3. Interface Design of Server .....	221
3.3.1. Structure and layout.....	221
3.3.2. Page content description.....	223
V. EVALUATION.....	235
1. TEST PLAN .....	235
1.1. Introduction.....	235
1.1.1. Purpose.....	235
1.1.2. Background information.....	235
1.1.3. Scope of testing .....	236
1.1.4. Constraints.....	237
1.1.5. Risk List.....	237
1.1.6. Training Needs .....	237
1.2. Requirement For Test.....	238

1.2.1. Test items.....	238
1.2.2. Acceptance test criteria .....	239
1.3. Test strategy .....	239
1.3.1. Test type.....	239
1.3.2 Test stages .....	241
1.4. Resource .....	242
1.4.1 Human resource.....	242
1.4.2 Environment .....	243
1.5. Test Milestones .....	243
1.6. Deliverables.....	244
2. TEST REPORT .....	245
2.1. Integration test .....	245
2.2. System test.....	245
3. CHECK LIST .....	246
VI. IMPLEMENTATION.....	248
1. DEVELOPMENT.....	248
2. PROGRESS REPORT.....	248
VII. INSTALLATION GUIDE AND USER MANUAL .....	251
1. INSTALLATION GUIDE .....	251
1.1. English Version.....	251
1.1.1. Prerequisites .....	251
1.1.2. Installation procedures.....	251
1.1.3. Uninstall procedures.....	251
1.2. Japanese Version .....	251
1. 1. 1. 前提条件 .....	251
1. 1. 2. インストール手順.....	251
1. 1. 3. アンインストール手順.....	252
2. USER MANUAL .....	252
2.1. Select a screen.....	252
2.2. Scan document .....	253
2.3. Select text on scanned document.....	253
2.4. Search word in Japanese dictionary.....	254
2.5. Search kanji in Kanji dictionary.....	255
2.6. View history .....	256
2.7. View all favorite words .....	256
2.8. Clear history.....	257
2.10. Select Dictionary package.....	258
2.11. Delete Dictionary package .....	259
2.12. Download Dictionary package .....	259
VII. CONCLUSION .....	261
1. LESSONS LEARN.....	261
2. FUTURE WORK.....	261
APPENDICES .....	262
1. SURVEY SUMMARY .....	262
2. MEETING MINUTES .....	264

3. CONFIGURATION MANAGEMENT.....	267
3.1. CI Identification and Naming Convention .....	267
3.2. Version Numbering Rules <sup>[6]</sup> .....	269
3.3. Directory Structure .....	270
4. CODING CONVENTION .....	271
4.1. Coding Convention.....	271
4.2. Comment Convention .....	272
REFERENCES:.....	273

## ACRONYMS & ABBREVIATIONS

Abbreviations	Description	Note
AKP	Anki Pan	
CI	Configuration Items	
OCR	Optical Character Recognition	
MM	Meeting Minute	
PM	Project Manager	
PMP	Project Management Plan	
SDD	Software Detail Design	
SRS	Software Requirement Specification	
SUM	Software User Manual	
W#	Week#	
WR	Weekly Report	
SAD	Software Architecture Design	
FRS	Functional Requirement Specification	
QA	Quality Assurance	
PM	Project Manager	
SRS	Software Requirement Specification	
TM	Team Member	
De	Designer	
DB	Database	
.db	SQLite Database file extension	
.ico	Icon file extension	
PK	Primary Key	
FK	Foreign Key	

## I. INTRODUCTION

### 1. PROJECT INFORMATION

- ❖ **Project name:** Anki Pan
- ❖ **Project code:** AKP
- ❖ **Project type:** Android application
- ❖ **Project category:** New development
- ❖ **Business domain:** Education
- ❖ **Project manager:** Bui Hong Nhung
- ❖ **Time line:** May 4<sup>th</sup> 2015 to August 15<sup>th</sup> 2015

### 2. PROJECT TEAM

- ❖ **Supervisor**

Name	Email	Title
Tran Binh Duong	duongtb@fpt.edu.vn	Supervisor

- ❖ **Team member**

Name	Student Roll	Email	Role
Bui Hong Nhung	SE02082	NhungbhSE02082@fpt.edu.vn	Leader
Vu Hong Thai	SE02654	thaivhSE02654@fpt.edu.vn	Member
Le Thanh Binh	SE02917	binhltSE02917@fpt.edu.vn	Member
Nguyen Duc Huy Hoang	SE02651	HoangndhSE02651@fpt.edu.vn	Member
Pham Trong Nghia	SE02781	NghiaptSE02781@fpt.edu.vn	Member
Ngo Thi Mai	SE02218	MaintSE02218@fpt.edu.vn	Member

### 3. INITIAL IDEAL

#### 3.1. Background

Japan is one of the most advanced countries in the world. In addition, Japanese culture is spreading very quickly to other countries through manga, anime,... Therefore, more and more people study Japanese. A research of JPF (Japan Foundation for Culture Exchange) shows that until 2012, there are 3,985,669 people studying Japanese, 63,805 Japanese teachers and 16,046 Japanese educational institutions. Those numbers are still constantly growing bigger and bigger.

表1-1-1 機関数・教師数・学習者数

	2012年	2009年
機関(機関)	16,046	14,925
教師(人)	63,805	49,803
学習者(人)	3,985,669	3,651,232

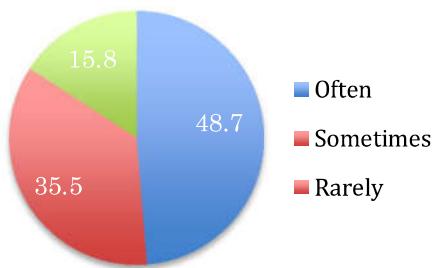
(機関 : Organization 教師 : Teacher 学習者 : Student)

Figure 1: JPF research in 2013

Most of people who study Japanese find it difficult to read Japanese document. Japanese includes a lot of Chinese-like characters called Kanji. Without knowing how to read the Kanji, it will be difficult to type those Kanji to search for their meaning.

We conducted a research among students who are studying Japanese. Those students often read books, comic, newspapers written in Japanese (see Chart 1). The result shows that 88.2% out of 78 students doing the survey find it difficult to read Japanese document (see Chart 2).

Do you often read Japanese document?



Is it hard to read Japanese?

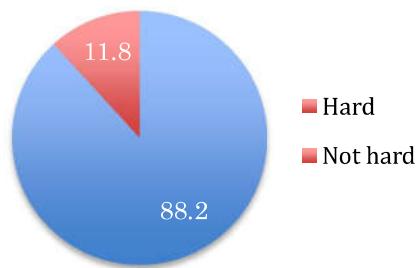


Chart 1

Chart 2

Almost of the students want an application that can scan Japanese in documents. More than 60% of them need help in translate the document, and search for the meaning of words.

(See AK\_SurveySummary in Appendix for detail about this survey)

## 3.2. Existing Products

### 3.2.1. Google translate<sup>[2]</sup>

Google translate is considered as the best application in language processing. It provides user a completed services including character recognition, speech to text and translate.

#### Functions:

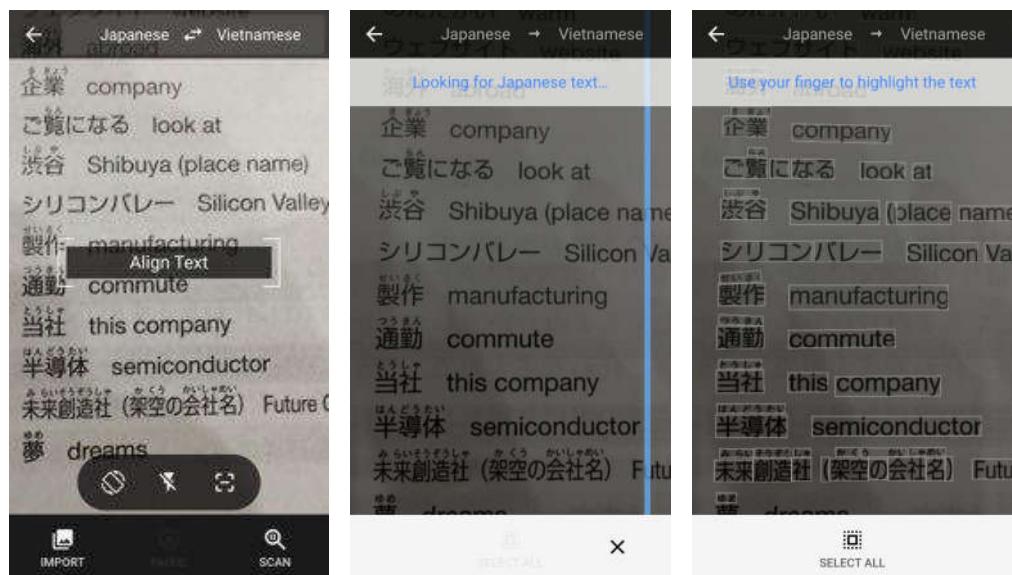
- Scan printed document and detect characters of many language, include Japanese
- Allow user to import image to detect characters
- Translate to many other language
- Convert speech to text and translate
- Save translated words and documents
- Bookmark words / documents
- Copy words / documents
- Turn on flash to take image

#### Some screenshots of Google translate:

Main screen



Scan and recognize text:



Translate to other languages:



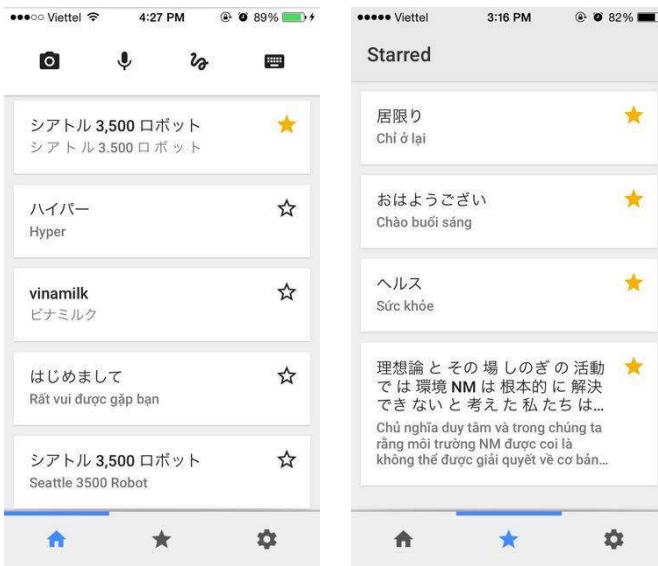
### Speech to text



Rất vui được gặp bạn →



### Save to history and bookmark

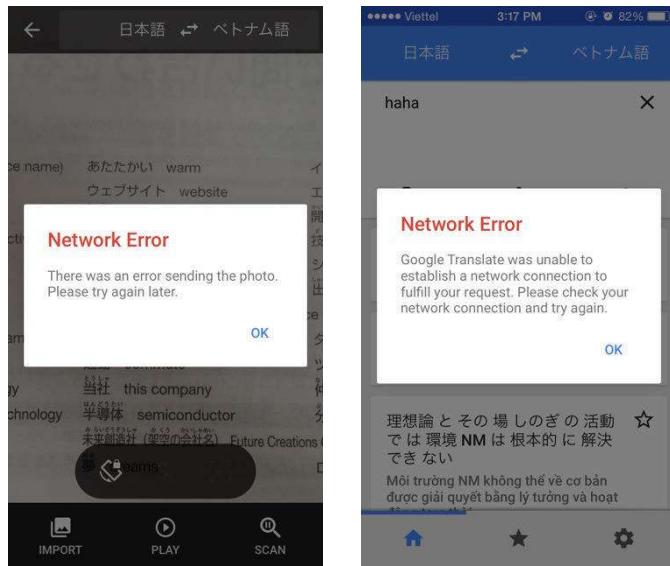


#### Pros:

- Interface design is simple and easy to use
- Document scanning is fast
- Character recognition is in high correction rate
- Recognize many language
- Translate from / to many language

#### Cons:

- Require Internet access for character recognition and translation



- Translation may return wrong meaning of words
- Does not provide other information about words (word-class, examples,...). Only the meaning of words may not satisfies people who study Japanese and want to understand the words deeply.

### **3.2.2. CamDictionary<sup>[3]</sup>**

CamDictionary is another good application in character recognition

#### **Functions:**

- Scan and recognize printed characters of many languages in three ways:
  - + User touch somewhere in the screen to show application the area to scan
  - + User move the application eye to the text area that they want to scan
  - + User capture to document, and then choose a part of the document to scan
- Allow user to import image to detect characters
- Turn on flash to take image
- Translate from / to many other languages
- Save translated words to History or My Vocabulary
- Download dictionary and search words in dictionary
- Send words and words' meaning through message or email

#### **Some screenshots of CamDictionary:**

## Scan and recognize text:



## Translate:



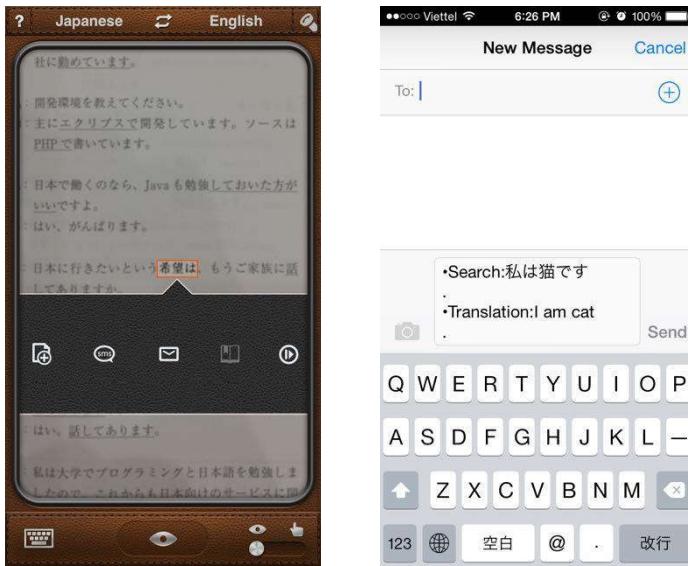
## Save to history or my vocabulary



## Download dictionary:



Send through email or message:



### Pros:

- The application is interactive. User can choose to scan document in 3 ways. The design for scanning screen is interesting.
- Text scan and character recognition is highly correct.
- The scanning function works offline
- User can purchase dictionary to search words, which give a more correct meaning and other information like word-class, example,...

### Cons:

- Require Internet access for translation
- User has to pay for the dictionaries

### 3.2.3. JDict<sup>[4]</sup>

Jdict is one of the most popular Japanese – Vietnamese dictionary application

#### Functions:

- Look up words in Japanese – Vietnamese dictionary, Vietnamese – Japanese dictionary and Kanji dictionary
- Save words to history
- Bookmark words

#### Some screenshots of JDICT

### Search words in dictionary



### Save to history and bookmark



#### Pros:

- Simple and easy-to-understand interface design
- Give detail meanings, examples, word-class, pronunciation of words
- Can search for about 700.000 Japanese words, and almost Kanji
- Explain words' meaning in Vietnamese, which is helpful for Vietnamese people.

#### Cons:

- User has to input words to search for the meaning. If they do not know how to read Kanji, they can not type the Kanji to search. It will take time to draw the Kanji from keyboard.

### 3.2. Our Proposal

We will develop a single application that support people who studying Japanese everything they need for reading printed Japanese document. The application will have the below main functions:

- Scan text in Japanese document
- Recognize Japanese characters
- Allow user to download free dictionaries
- Search for meaning of words, Kanji in dictionary
- Save words to history for user's later review
- Add words to favorite list

The application will be interactive, easy to use and work offline.



*Figure 2: Doraemon's copying toast*

We name the application Anki Pan, which means copying toast in English. Anki Pan is a Doraemon's gadget which allows user to memorize any contents 'printed' on it. To use the copying toast, one should press one side of the toast on a book page. The text on the book then will be printed on the toast. Eating the toast later, eater can remember everything that has been printed on the toast. Just like how to use the copying toast, user using our application will take their phone above the book page and press the capture button. The book page with text on it then will be saved (printed) to the application. The application will extract text from the book page so that user can require application to translate or to look up meaning of words. By doing this, user can understand the content of the book, remember words and their meaning, just like swallowing the content of book by eating a toast.

Our application in compare to Google Translate, CamDictionary and JDICT:

Function	Google Translate	CamDictionary	JDict	Anki Pan
Scan and recognize text	Many language	Many language	✗	Japanese only
Translate	✓	✓	✗	✗
Search words in dictionary	✗	✓	✓	✓
Free Dictionary	✗	✗	✓	✓
Save words to history	✓	✓	✓	✓
Bookmark words	✓	✓	✓	✓
Work offline	✗	✗	✓	✓