



# Handover and completion report

Extraction of Structured  
Information from Dictionary PDFs

Caleb Smith  
Josephine Paculio  
Joshua Dib  
Roger Bernardo

## Executive Summary

In response to the tedious process transforming a default dictionary file into a standard electronic dataset an automatic alternative has been developed. The implemented solution for this task is a local web application which allows the user to select a file, display the output, manipulate data, and export the collection.

The structure will first provide the reader with a final overview of the system and an insight into the variations from the original proposal such as the screen designs, and new features that have been developed as the scope grew. In addition, extensive testing will be shown to ensure the system is a complete product followed by outstanding issues the enhance the experience and left to the discretion of the end user to decide how to continue,

This report is paramount to all parties as it provides capstone to the system and finalise all the work that has been completed and the potential areas in which the product can grow.

## Contents

Executive Summary .....	1
1. Introduction .....	3
2. Test Results .....	3
3. Variations Report .....	12
3.1 Screen Design Variation .....	12
3.2 Use Case Variation .....	13
4. Known Bugs .....	15
5. Outstanding development work/Further improvements .....	19
6. Conclusion.....	19

# 1. Introduction

A problem in the data analysis and research field had been identified, that being the long and arduous task of converting files into a dataset, in response a solution has been developed to automate this process. This document illustrates how the project has evolved from the initial proposal and lists the changes the design, data, and functions have undergone. Moreover, it will describe any loose ends that can then be followed up or left depending on the desired scope for the future.

Following on from this document the reader will be able to close the project knowing it has been completed and identifies future possibilities.

# 2. Test Results

Use Case/ Feature	Attach an OCR'd pdf file and extract text
Test purpose	To check the accuracy of the program
Expected Results	See Individual Test Data Sets
Success/Failure	Success

Test 1					
		Total number of entries/keywords	Actual number of entries/keywords collected	Error (keyword inside a keyword)	Error (Incorrect keyword/s)
Attached PDF File Name (Mandatory)	djabugay_dictionary_quinn et al.pdf	1085	1085	0	0
Start page (Mandatory)	8				
End page (Mandatory)	22				
Keyword (Mandatory)	axe				
Header	None				
Footer	None				

Test 2				
	Total number of entries/keywords	Actual number of entries/keywords collected	Error (keyword inside a	Error (Incorrect keyword/s)

				keyword )	
<b>Attached PDF File Name (Mandator y)</b>	harvey_larrakia-dictionary2004.pdf	4372	4372	0	0
<b>Start page (Mandator y)</b>	2				
<b>End page (Mandator y)</b>	91				
<b>Keyword (Mandator y)</b>	dadugu-da				
<b>Header</b>	None				
<b>Footer</b>	None				

Test 3					
		Total number of entries/keywords	Actual number of entries/keywords collected	Error (keyword inside a keyword)	Error (Incorrect keyword/s)
<b>Attached PDF File Name (Mandator y)</b>	lillie_girawa[1999].pdf	2994 (according to pdf doc)	2997	0	3
<b>Start page (Mandator y)</b>	1				
<b>End page (Mandator y)</b>	47				
<b>Keyword (Mandator y)</b>	aipar				
<b>Header</b>	-a				
<b>Footer</b>	4/27/2005				

Test 4					
		Total number of entries/keywords	Actual number of entries/keywords collected	Error (key word inside a keyword)	Error (Incorr ect keywo rd/s)
<b>Attach ed PDF File Name</b>	PamaNyungan_Yolngu_Djinang_waters_djinangviv1983.pdf	3363	3363	0	0

<b>(Mandatory)</b>					
<b>Start page (Mandatory)</b>	1				
<b>End page (Mandatory)</b>	135				
<b>Keyword (Mandatory)</b>	baba				
<b>Header</b>	None				
<b>Footer</b>	None				

<b>Test 5</b>					
		Total number of entries/keywords	Actual number of entries/keywords collected	Error (keyword inside a keyword)	Error (Incorrect keyword/s)
<b>Attached PDF File Name (Mandatory)</b>	lowe_gupapyungu2004.pdf	4120 (according to pdf doc)	4115	0	5
<b>Start page (Mandatory)</b>	11				
<b>End page (Mandatory)</b>	199				
<b>Keyword (Mandatory)</b>	babalamirri				
<b>Header</b>	bagapaga				
<b>Footer</b>	Yolngu - English Dictionary B. Lowe © ARDS Inc. September 2004.				

<b>Test 6</b>					
		Total number of entries/keywords	Actual number of entries/keywords collected	Error (keyword inside a keyword)	Error (Incorrect keyword/s)
<b>Attached PDF File Name (Mandatory)</b>	Mead - 2018 - A brief dictionary of Bobongko (Central Sulawesi).pdf	3410	3724	314	0
<b>Start page (Mandatory)</b>	4				
<b>End page</b>	57				

<b>(Mandatory)</b>					
<b>Keyword (Mandatory)</b>	ala				
<b>Header</b>	agu				
<b>Footer</b>	3				

#### Test 7

		Total number of entries/keywords	Actual number of entries/keywords collected	Error (keyword inside a keyword)	Error (Incorrect keyword/s)
<b>Attached PDF File Name (Mandatory)</b>	Wirangu - English Dictionary.pdf	1552	1552	0	0
<b>Start page (Mandatory)</b>	0				
<b>End page (Mandatory)</b>	84				
<b>Keyword (Mandatory)</b>	abubaldha				
<b>Header</b>	Draft only – 19 March 2010				
<b>Footer</b>	1				

#### Test 8

<b>****Pages not aligned and have different orientation</b>		Total number of entries/keywords	Actual number of entries/keywords collected	Error (keyword inside a keyword)	Error (Incorrect keyword/s)
<b>Attached PDF File Name (Mandatory)</b>	whitby_dadibi1990_o.pdf	N/A	N/A	N/A	N/A
<b>Start page (Mandatory)</b>	4				
<b>End page (Mandatory)</b>	50				
<b>Keyword (Mandatory)</b>	abe				
<b>Header</b>	Dadibi, Tok PMn, English [dogonel				
<b>Footer</b>	Yabe 5				

#### Test 9

		Total number of entries/keywords	Actual number of entries/keywords collected	Error (keyword inside a keyword)	Error (Incorrect keyword/s)
<b>Attached PDF File Name (Mandatory)</b>	hooley_buang2006_s.pdf	3970	3970	0	0
<b>Start page (Mandatory)</b>	17				
<b>End page (Mandatory)</b>	406				
<b>Keyword (Mandatory)</b>	aba				
<b>Header</b>	Page 18				
<b>Footer</b>	11/3/2006				

#### Test 10

****Pages and words are not aligned		Total number of entries/keywords	Actual number of entries/keywords collected	Error (keyword inside a keyword)	Error (Incorrect keyword/s)
<b>Attached PDF File Name (Mandatory)</b>	Hersri - 1982 - A Language from Buru Wayapo-Indonesian Wordlist.pdf	N/A	N/A	N/A	N/A
<b>Start page (Mandatory)</b>	4				
<b>End page (Mandatory)</b>	23				
<b>Keyword (Mandatory)</b>	abat				
<b>Header</b>	135				
<b>Footer</b>	None				

#### Test 11

		Total number of entries/keywords	Actual number of entries/keywords collected	Error (keyword inside a keyword)	Error (Incorrect keyword/s)
<b>Attached PDF File Name (Mandatory)</b>	Nyangumart a Dictionary Wordlists Cover 2008 Bookmarked Burgman.pdf	2747	2747	0	0
<b>Start page (Mandatory)</b>	58				



<b>End page (Mandatory)</b>	150				
<b>Keyword (Mandatory)</b>	-arrangu				
<b>Header</b>	-ja				
<b>Footer</b>	Nyangumarta – English Dictionary 2008				

#### Test 12

<b>****Keywords have inconsistent indentation</b>		Total number of entries/keywords	Actual number of entries/keywords collected	Error (keyword inside a keyword)	Error (Incorrect keyword/s)
<b>Attached PDF File Name (Mandatory)</b>	Evans - 2003 - Kamus Kaili-Ledo - Indonesia - Inggris (Kaili-Ledo - Indonesian - English dictionary).pdf	N/A	N/A	N/A	N/A
<b>Start page (Mandatory)</b>	15				
<b>End page (Mandatory)</b>	301				
<b>Keyword (Mandatory)</b>	agaiana				
<b>Header</b>	aa				
<b>Footer</b>	1				

#### Test 13

<b>****Keywords have inconsistent indentation</b>		Total number of entries/keywords	Actual number of entries/keywords collected	Error (keyword inside a keyword)	Error (Incorrect keyword/s)
<b>Attached PDF File Name (Mandatory)</b>	Evans - 2003 - Kamus Kaili-Ledo - Indonesia - Inggris (Kaili-Ledo - Indonesian - English dictionary).pdf	N/A	N/A	N/A	N/A

<b>Start page (Mandatory)</b>	305				
<b>End page (Mandatory)</b>	395				
<b>Keyword (Mandatory)</b>	abu-abu				
<b>Header</b>	absen				
<b>Footer</b>	291				

#### Test 14

<b>****Keywords have inconsistent indentation</b>		Total number of entries/keywords	Actual number of entries/keywords collected	Error (keyword inside a keyword)	Error (Incorrect keyword/s)
<b>Attached PDF File Name (Mandatory)</b>	Evans - 2003 - Kamus Kaili-Ledo - Indonesia - Inggris (Kaili-Ledo - Indonesian - English dictionary).pdf	N/A	N/A	N/A	N/A
<b>Start page (Mandatory)</b>	399				
<b>End page (Mandatory)</b>	495				
<b>Keyword (Mandatory)</b>	able				
<b>Header</b>	abandoned				
<b>Footer</b>	385				

#### Test 15

<b>****Column 2 is read first before column 1</b>		Total number of entries/keywords	Actual number of entries/keywords collected	Error (keyword inside a keyword)	Error (Incorrect keyword/s)
<b>Attached PDF File Name (Mandatory)</b>	KaytDictionary_proof3.pdf	N/A	N/A	N/A	N/A
<b>Start page (Mandatory)</b>	38				
<b>End page (Mandatory)</b>	655				
<b>Keyword</b>	ahe				

<b>(Mandatory)</b>					
<b>Header</b>	aahlwerrnge				
<b>Footer</b>	SEE PAGE 8 FOR AN EXPLANATION OF THE ENTRY LAYOUT				

Test 16					
<b>****Different format than other pdf files</b>		Total number of entries/keywords	Actual number of entries/keywords collected	Error (keyword inside a keyword)	Error (Incorrect keyword/s)
<b>Attached PDF File Name (Mandatory)</b>	verstraete-rigsby_yintyingka2015v2.pdf	N/A	N/A	N/A	N/A
<b>Start page (Mandatory)</b>					
<b>End page (Mandatory)</b>					
<b>Keyword (Mandatory)</b>					
<b>Header</b>					
<b>Footer</b>					

Test 17					
<b>****Keywords have inconsistent indentation</b>		Total number of entries/keywords	Actual number of entries/keywords collected	Error (keyword inside a keyword)	Error (Incorrect keyword/s)
<b>Attached PDF File Name (Mandatory)</b>	Anceaux - 1987 - Wolio Dictionary (Wolio-English-Indonesian) Kamus Bahasa Wolio (Wolio-Inggeris-Indonesia).pdf	N/A	N/A	N/A	N/A
<b>Start page (Mandatory)</b>	21				
<b>End page (Mandatory)</b>	215				
<b>Keyword (Mandatory)</b>	aa				

<b>Header</b>	agori				
<b>Footer</b>	none				

Test 18					
<b>****Pages not aligned</b>		Total number of entries/keywords	Actual number of entries/keywords collected	Error (keyword inside a keyword)	Error (Incorrect keyword/s)
<b>Attached PDF File Name (Mandatory)</b>	Heath - 1982 - Nunggubuy Dictionary.pdf	N/A	N/A	N/A	N/A
<b>Start page (Mandatory)</b>	8				
<b>End page (Mandatory)</b>	162				
<b>Keyword (Mandatory)</b>	ba-dhawiwi-j				
<b>Header</b>	none				
<b>Footer</b>	1				

<b>Use Case/Feature</b>	Form Validation
<b>Test Purpose</b>	To ensure all implemented validation works as intended
<b>Expected Results</b>	Test individual data
<b>Success/Failure</b>	Success

#### Test Results

Form Label	Set 1	Set 2	Set 3
Attach a file	Attach PDF	PDF not attached	
Enter Page Numbers	2 to 80	309320932093	QWKOQ
Add a keyword	Burrara	Z	393232
Footer or header	Checkbox yes/no		
Expected Result	Pass	Fail	
Results	Success	Fail	Fail

## 3. Variations Report

### 3.1 Screen Design Variation

#### Form Input

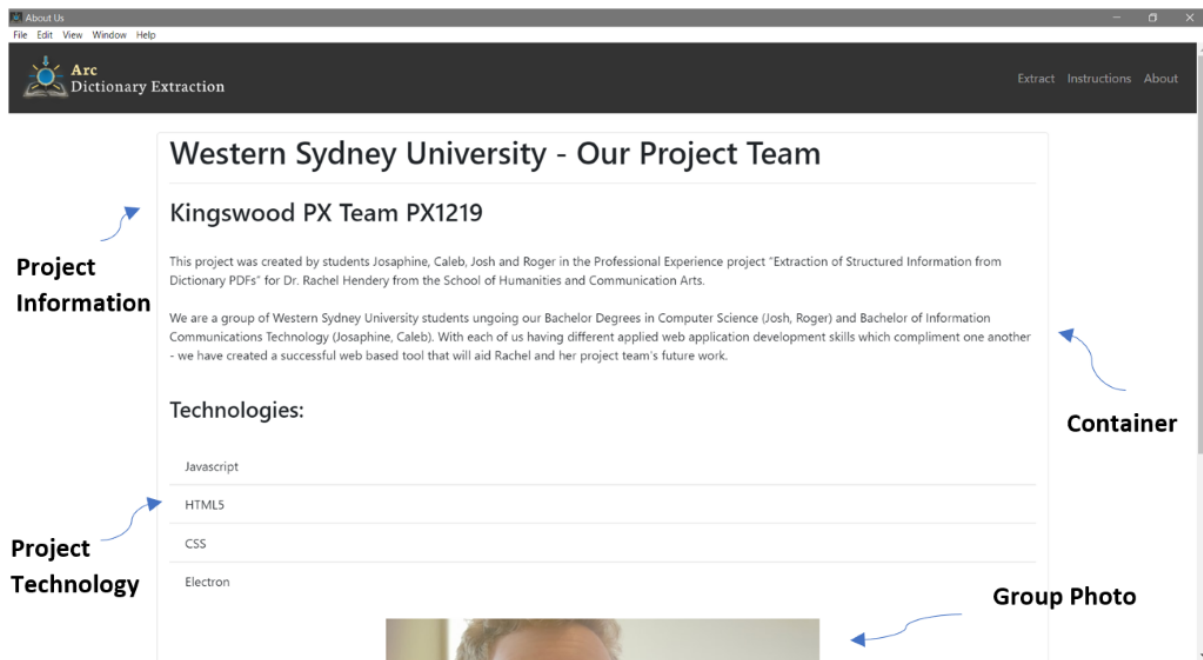
The screenshot shows the 'Arc Dictionary Extraction' application window. The interface includes a menu bar (File, Edit, View, Window, Help), a toolbar with 'Extract', 'Instructions', and 'About', and a main content area titled 'Dictionary Query Fields'. The form contains several input fields: a file path ('Hersri - 1982 - A Language from Buru Wayapo-Indonesian Wordlist.pdf'), a 'Browse' button, a 'Please enter the page number where you want to start the search:' field (value: 1), a 'Please enter the page number where you want to end the search:' field (value: 3), a 'Please enter a keyword here:' field (value: Abubaldha), and two checkboxes for 'Does the PDF file have a Header/s' and 'Does the PDF file have a footer/s'. An 'Extract' button is at the bottom. Annotations point to various elements: 'Product Logo' (Arc Dictionary Extraction logo), 'Input Validation' (red border around the keyword field), 'Application Footer' (Developed 2019 Joroca), 'Navigation' (Extract, Instructions, About links), 'Collapsible Card' (the main form area), and 'Extraction Form' (the form fields themselves).

#### Data Output

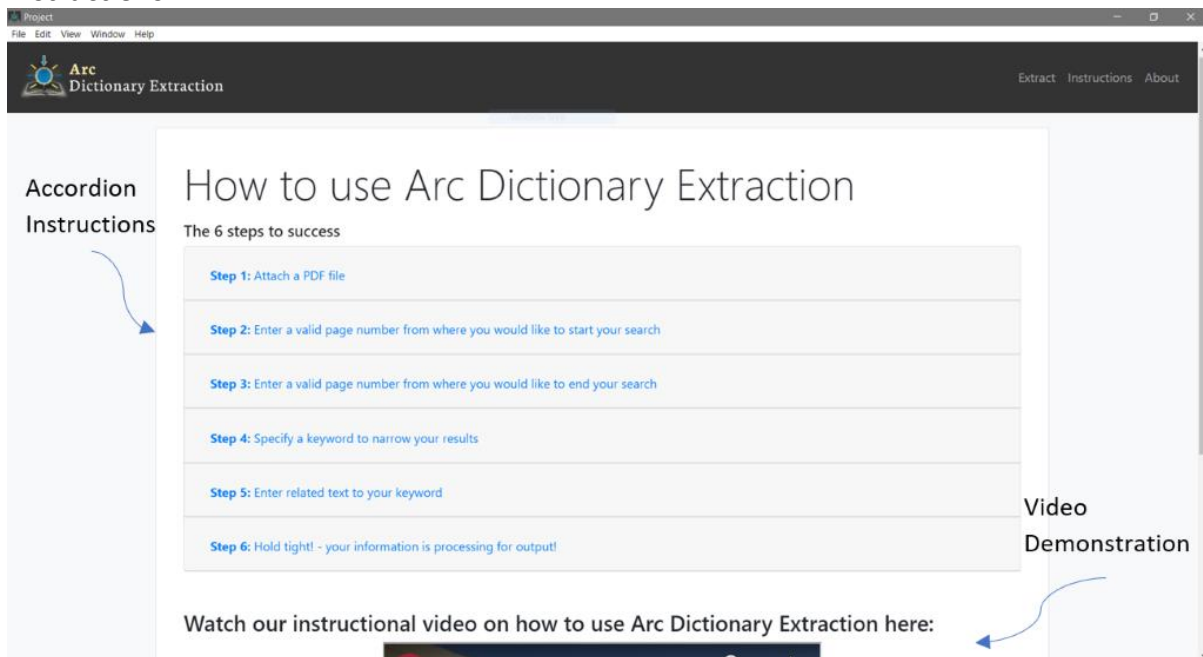
The screenshot shows the 'Arc Dictionary Extraction' application window displaying a table of extracted data. The table has four columns: 'Process Number', 'Keyword', 'Part of Speech', and 'Related Text'. The table is filtered to show entries 11 to 20 of 64. Annotations point to various elements: 'Application Menu' (File, Edit, View, Window, Help), 'Table options' (Add, Edit, Delete, CSV, PDF), 'Data Filter' (Search field), 'Table Information' (Showing 11 to 20 of 64 entries), 'Data Output' (the table itself), and 'Pagination' (Previous, 1, 2, 3, 4, 5, 6, 7, Next).

Process Number	Keyword	Part of Speech	Related Text
10	anak	n.	1) female animal, 2) his mother, jo anak big net bag
11	anaktei	n.	PaVoSi : parent's younger sister. [Note: Tok Pisin "liklik mama"]
12	anai	n.	taro (type).
13	anastein	vr.	they looked at each other.
14	ananim	v.	I see : I look. Syn: aparom.
15	ansu	adj.	not ripe.
16	ani	n.	(my) mother. anin; (your) mother. anak; (his) mother. aninok; (our) mother. aninaka; (your/their) mother. aninawa; mother(of you-2,they-2). anikot; mothers in general.
17	anikas	n.	juice(of taro).
18	anor	n.	bamboo used for carrying water.
19	anour	n.	Two-leaf tree. Syn: anur/onur.

#### About us



## Instructions

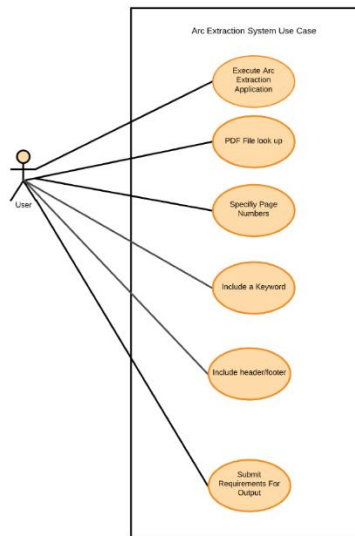


## 3.2 Use Case Variation

Type	Variation	Initiated by	Implemented
Functional	Add a new use case - Added validation to validate the form	Supervisor/Team	Partly implemented – further development required in following week.

<b>Functional</b>	Added a header as part of form output	Team	Implemented
<b>Functional</b>	Added a footer as part of form output	Team	Implemented
<b>Functional</b>	Changes in platform to Electron desktop application	Supervisor/Team	Implemented
<b>Functional</b>	Navigation addition	Team	Implemented

### Updated System Use Case



## 4. Known Bugs

### Bug 1 – Keywords inside a keyword

- Related text of a keyword that also have an indented related text are classified as a keyword instead of a related text of the original keyword.

A - a	
<p>agu' ash. abu. [repl. Adu. agu]</p> <p>agu'on ashy, covered in ashes. berabu. [repl. Adu. 'gray']</p> <p>ajaran (-&lt;Iv) horse. kuda.</p> <p>aje chin. dagu.</p> <p>akar (-&lt;MI)</p> <p>mongakari deceive. menipu.</p> <p>akol aren palm. sugarpalm. pohon aren.</p> <p>aku first person singular independent pronoun.</p> <p>pronomina persona pertama tunggal bebas.</p> <p>ala</p> <p>mongala take. mengambil.</p> <p>moko'ala able to. dapat. sanggup.</p> <p>aladi' taro. talas. keladi.</p> <p>alangan (-&lt;MI) hinderance, obstacle. halangan.</p> <p>ali-ali (-&lt;GM) younger sibling. adik.</p> <p>alibobos chicken louse. tungau. kutu kecil</p> <p>sekali yang terdapat pada ayam.</p> <p>alin</p> <p>ma'alin clean. bersih.</p> <p>mompoko'alin make clean. membersihkan.</p> <p>mongalini trim, slash off (as side branches).</p> <p>menetak. membacok.</p> <p>alus (-&lt;MI)</p> <p>mompoko'alus crush (pound, grind, etc.) into fine particles. menghaluskan.</p> <p>ambang</p> <p>ma'ambang ashamed. malu.</p> <p>aminsing cheek. pipi.</p> <p>ampas mat. tikar.</p> <p>ampung (-&lt;MI) forgiveness. ampun.</p> <p>mongampungi forgive. mengampunkan, memaafkan.</p> <p>anak child. anak.</p> <p>monganak bear, give birth. melahirkan.</p> <p>monga'anakakon beget. memperanakan.</p> <p>poku'anakon niece, nephew. kemanakan.</p> <p>[repl. Adu. pokuanakon]</p>	<p>anak itu a child with only one living parent.</p> <p>anak yang bapaknya atau ibunya sudah meninggal. See: unon.</p> <p>anak lolong baby, infant. bayi.</p> <p>anak mantu child-in-law. menantu. See: monugang.</p> <p>anak potuu' adopted child. anak pungut.</p> <p>See: tuu'.</p> <p>anak rapi' twins. anak kembar.</p> <p>anak tampalui' last born child. anak bungsu.</p> <p>anak uluttuong first born child. anak sulung.</p> <p>ande' monkey. monyet.</p> <p>ane' termite. rayap. anai-anai.</p> <p>ansang fish gills. insang.</p> <p>antimun cucumber. ketimun.</p> <p>anu whatsit. anu.</p> <p>anyam</p> <p>mongananyam weave, plait. menganyam.</p> <p>apu fire. api.</p> <p>ara third person plural independent pronoun.</p> <p>pronomina persona ketiga jamak. [repl. Adu. ara third person (singular and plural) pronoun]</p> <p>aro</p> <p>umaro shout. berteriak.</p> <p>arus</p> <p>ma'arus smooth. halus, licin.</p> <p>asi' also, too, juga.</p> <p>asu' dog. anjing. [repl. Adu. asu]</p> <p>ata'</p> <p>pongata' hope (n). harapan.</p> <p>ate liver. hati.</p> <p>ate bura' lungs. paru-paru.</p> <p>atop roof, roofing thatch. atap.</p> <p>atu hundred. ratus. [repl. Adu. gatut]</p>

anak	child. anak. monganak bear, give birth. melahirkan. monga'anakakon beget. memperanakan. poku'anakon niece, nephew. kemanakan. [repl. Adu. pokuanakon]	17
anak	ilu a child with only one living parent. anak yang bapaknya atau ibunya sudah meninggal. See: unon.	18
anak	lolong baby, infant. bayi.	19
anak	mantu child-in-law. menantu. See: monugang.	20
anak	potuu' adopted child. anak pungut. See: tuu'.	21
anak	rapi' twins. anak kembar.	22
anak	tampalui' last born child. anak bungsu.	23
anak	uluttuong first born child. anak sulung.	24



## Bug 2 – Special characters

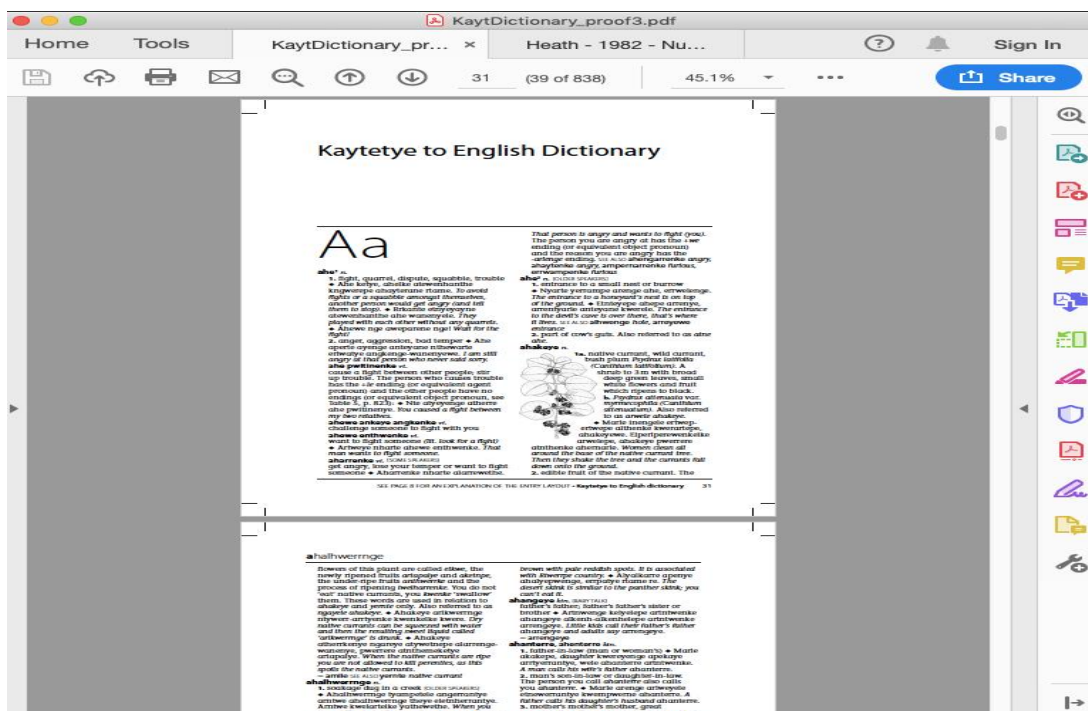
- When the program encounters a special character, it will ignore all the remaining text including the special character within the same line.

ajaran (<Jv) horse. kuda.  
aje chin. dagu.  
akar (<MI)  
mongakari deceive. manipu.  
akal

ajaran	(	1
aje	chin. dagu.	2
akar	(mongakari deceive. manipu.	3

## Bug 3 – Column 2 is read first before column 1

- The program will not have correct results if the pdf file's column 2 is read first before column 1. The program will not detect if column 2 is read first before column 1.



### Bug 4 – Different page sizes and orientation

- The program will not detect if the pdf file has different page sizes and orientation. The program will not have accurate results as it depends heavily on difference in indentation between the keyword and related text.

[illegible]

## Bug 5 – Header and footer

- The program will treat the "A" in the first page as the header since it is the highest text/string. Because "A" is treated as the header, its y-coordinate will be used as a starting point to trim the remaining text above it. This will then result in an inaccurate result because some of the keywords and related text in the second page will be deleted because it got the same or higher y-coordinate as the "A".

# A

**a-** verbal prefix: actor third person singular and plural/ awalan kk untuk pelaku persona ketiga tunggal dan jamak. (WL.25)

**-a** I verbal suffix: subjunctive/ akhiran kk.: semoga, mudah-mudahan. *mataua ingkoo* – that you may know/ mudah-mudahan engkau tahu!

**II** verbal suffix: object third person singular and plural/ akhiran kk: obyek persona ketiga tunggal dan jamak. (WL. 27)

**aa** interjection expressing disbelief: you don't mean to say so! what makes you think so? I can't believe it/ tidak mungkin! bagaimana bisa? saya tidak percaya!

**'aabhidi** (A) ks pious, devout/ saleh, alim.

**'aadhili** (A) ks just, righteous/ adil, berbudi.

**agoria** soon, before long/ segera, tak lama lagi. brh. *agori, agoagori, kaagoago, maago.*

**aaherati** (A) kb hereafter, world to come/ akhirat, alam baka.

**aahiri** (A) 1. kb end, conclusion/ akhir, penutup. 2. last, final/ terakhir, penghabisan.

**aii** I don't know/ saya tidak tahu! *i apai ali-angka o andimu, Aii? Aii!* – where has your brother gone, Aii? I don't know! Adikmu pergi ke mana, Aii? Saya tidak tahu!

**'aakili** (A) ks having the full use of one's reasons, of sound mind, sensible, responsible for one's actions/ akal, berakal, bertanggung jawab atas perbuatannya.

**'aalimu** (A) kb Moslem theologist, religious scholar/ orang pandai dalam hal agama Islam. *'aalimu waga'a* – pious scholar/ ahli agama yg saleh. Jm. *'ulamaa.*

**'aairifu** (A) wise person, sensible, intelligent/ orang bijaksana, arif, berpikiran sehat. *'aairifu kaamili* – wise and pious man/ orang yg bijaksana dan saleh.

**aba** I kb k.o. fish/ sb. ikan.

**II** kkt ask (a question)/ menanyakan. brh. *poabaaka, poabaki, poabakia.*

**abanka** kkt ask about, inquire after/ menanyakan. brh. (s.d.a.).

**abaki** kkt ask (somebody), ask from/ bertanya kepada (seorang). brh. (s.d.a.).

**abha** (A) father (term of address), daddy/ bapak (panggilan), ayah.

**abhujaadhi** (A) alphabet/ abjad, alfabet.

**abi** kkt regard as, treat as/ menganggap (sebagai), memperlakukan (sebagai). *aabiaku kantilena* – they regard me as their slave/ mereka menganggap saya sebagai budaknya (expression of modesty/ pernyataan kerendahan hati).

*saro abina* – surname, nickname/ nama panggilan.

**ada** kkt borrow/ meminjam. *ada i* – borrow from/ meminjam dari. *Kuada dhoi i amaku* – I borrow money from my father/ saya meminjam uang dari ayah saya. brh. *paada, pandaaka, poada, poadaada.*

**adari** kkt teach, train, break in (a horse)/ mengajar, melatih, mendidik. brh. *kaadari, pengkaadari, pengkaadartaka.*

**ade** kb chin/ dagu. *ade bingku* – prominent chin/ dagu yg menonjol.

**'adhabhu** (A) kb punishment by superior order, punishment from God or from the sultan, supernatural punishment/ penyiksaan dari Allah atau dari Sultan, penyiksaan gaib.

**'adhati** (A) kb adat, custom, customary law, manners/ adat, adat-istiadat, kebiasaan, takrama. *matas 'adhati* – be well-mannered, well-educated/ sopan, tahu sopan santun. *inda matas 'adhati* – ill-mannered, impertinent/ kurang ajar, tidak sopan. brh. *ko'a-dhati.*

**adhedhe** exclamation of pain or sorrow: owl ouch! ah! alas! seruan sebab sakit atau sedih: aduh! aih! ah! brh. *dhedhe*.

**adi** particle of interrogation: exactly, ever/ kata penambah dalam kalimat tanya: *adi tua apa?* – what ever is the situation? what exactly is the matter?/ bagaimana gerangan itu?

**ae** kb leg, foot, paw, lower course (of a river), river-mouth/ kaki, cakar, hilir, muara. *Kauna ae* – toe/ jari kaki. *randana ae* – foot-sole/ tapak kaki. brh. *koae.*

**aenti** (A) kb Koran verse/ ayat Quran.

**'aebhu** (A) kb blame, reflection, defect, shortcoming, disgrace/ kesalahan, cacat, aib, malu, arang di muka, celaan. brh. *i'aebhuaka, ka'aebhu, pa'aebhu.*

**'aebhuaka** kkt put to shame/ memalukan. brh. (s.d.a.).

**acilolji** (MI) kb watch/ arloji. bdk. *arilolji, aroloji.*

**agama** (MI) kb religion/ agama.

**ago** kkt help, save/ menolong, membantu, menyelamatkan, mencegah. brh. *kaago.*

**agoagori** kkt hasten, speed up, accelerate/ mempercepat, mencepatkan, melajukan.

brh. *agori, kaagoago, maago.* bdk. *aagoria.*

**agoi** kkt steal, rob of, purloin, pinch, pilfer, appropriate dishonestly, misappropriate, take by force/ mencuri, merebut, merampas, merampok, mengambil dengan kekerasan

agori

2

senjata, menggelapkan. *agoi parintana baluwu o peropa* – take by force the highest power in the state, commit a coup d'état/ merebut kekuasaan negara dengan kekerasan, melakukan kudeta. brh. *mangaagoi, poagoi.*  
**agori** kkt do hastily, do speedily, speed up/ melajukan, mempercepat, bikin lekas-lekas.

mu: istilah penunjuk untuk orang berpangkat tinggi yg lebih tua daripada yg ditegur. *atumpusku akamui Bontona Baluwu* – I was ordered by the honorable B.B./ saya disuruh Yang Mulia B.B. *andimui akamui* – his highness (used when it is not certain that the person referred to is older or younger than the

## 5. Outstanding development work/Further improvements

Due to time constraints, certain features have not been fully developed and have not been included in the final version of the project.

- Performing an extraction on PDF files that have not been read through OCR. At present the program is only able to extract the content of files which have been OCR-ed.
- Multiple file upload and extracting from multiple PDF files. This has been supplemented by the addition of tabs
- Saving the current state of the program. This feature will allow the user to save and resume existing table results.

## 6. Conclusion

Upon concluding this document there are several key takeaways including an understanding of how the project has evolved from the early stages of development, the testing of the system, and remaining tasks for the future.

Following on the end user will be provided with an end product and resources to continue progress in the future.