

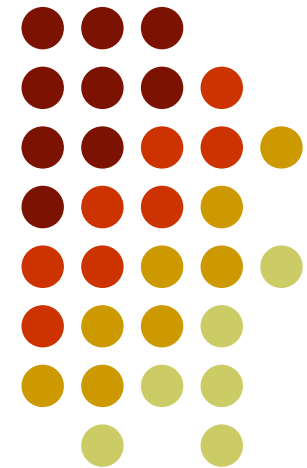
GlyphWiki: Shareable Glyph Database System for Chinese Characters

Koichi Kamichi

上地 宏一

Tokyo University of Technology

東京工科大学





What is GlyphWiki?

- online database for *Hanzi* glyphs (漢字字形)
- with web user interface
- typeface : *Song-ti* (宋体, 明朝体) glyphs
- data can use freely
- using KAGE system and WikiWikiWeb
- manage glyphs one by one



Background

- information processing for *Hanzi* texts
 - character coded sets
 - 2004 Japanese Ind. Std. : 10,050 *Hanzi*
 - 2003 Intl. Std. (≡Unicode) : 42,711 *Hanzi*
 - demand for using Variants which unified into other
 - large-scale *Hanzi* fonts (in Japan)
 - Mojikyo fonts : 100,000 *Hanzi*, partly non-free
 - GT fonts : 70,000 *Hanzi*, free in education
 - demand for completely free *Hanzi* fonts



Purpose of GlyphWiki

- collect all *Hanzi* glyphs / components
 - existing now / existed in the past
 - all of Variants 異体
 - create newly by user

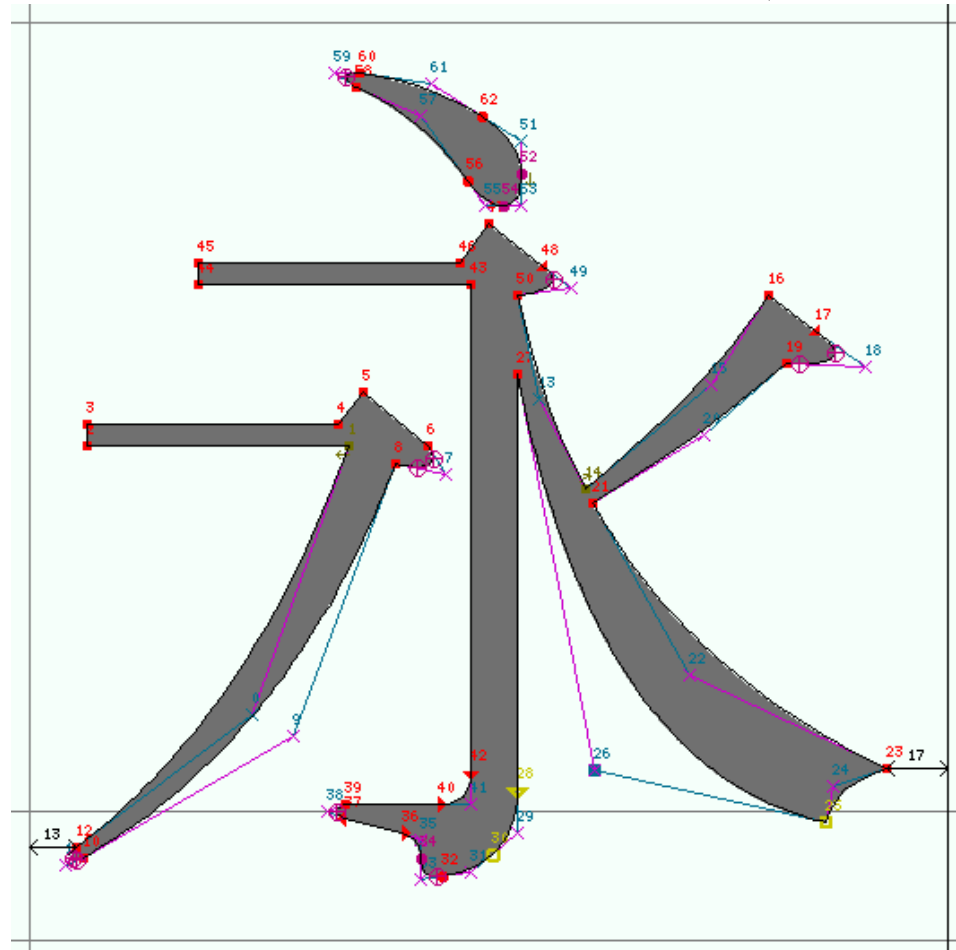
AND

- offer it to everybody
 - data type : glyph image (png, eps) / font (ttf)
- realization of large-scale environment of processing *Hanzi* on PCs



Using KAGE system (1)

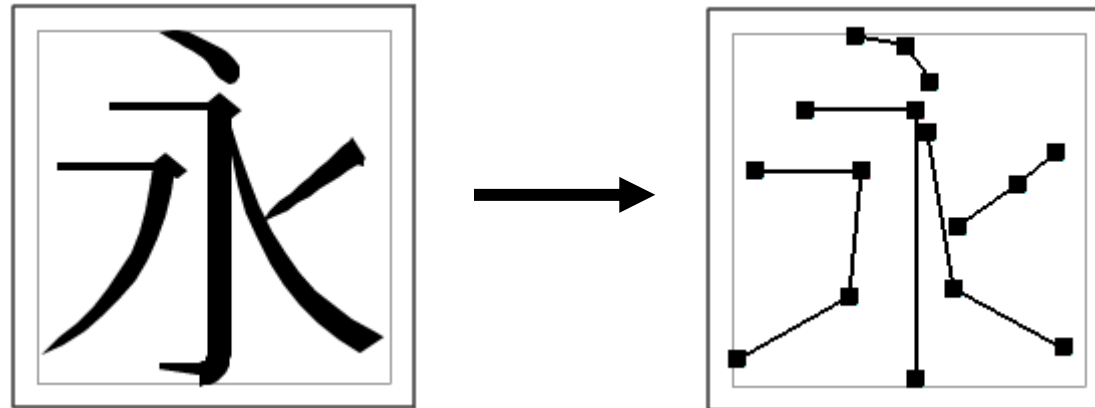
- ordinary : TrueType
 - general font format
 - draw outline of glyph
 - extremely difficult to design for non-designers





Using KAGE system (2)

- KAGE system
 - describe glyph by skeleton of each strokes
 - straight line / curve line / combination
 - 2 - 4 coordinates for each strokes
 - + type of shapes at the end of each strokes

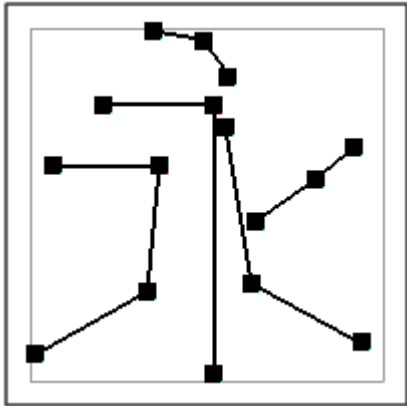


Using KAGE system (3)

example : 永



●	2:	7:	8:	73:	13:	98:	18:	110:	36
●	1:	0:	2:	48:	50:	103:	50		
●	1:	22:	4:	103:	50:	103:	184		
●	1:	0:	2:	22:	80:	76:	80		
●	2:	22:	7:	76:	80:	70:	143:	14:	174
●	2:	0:	7:	173:	71:	154:	87:	124:	108
●	2:	7:	0:	108:	57:	122:	139:	177:	168



↑ stroke type

↑ shape of top of the stroke

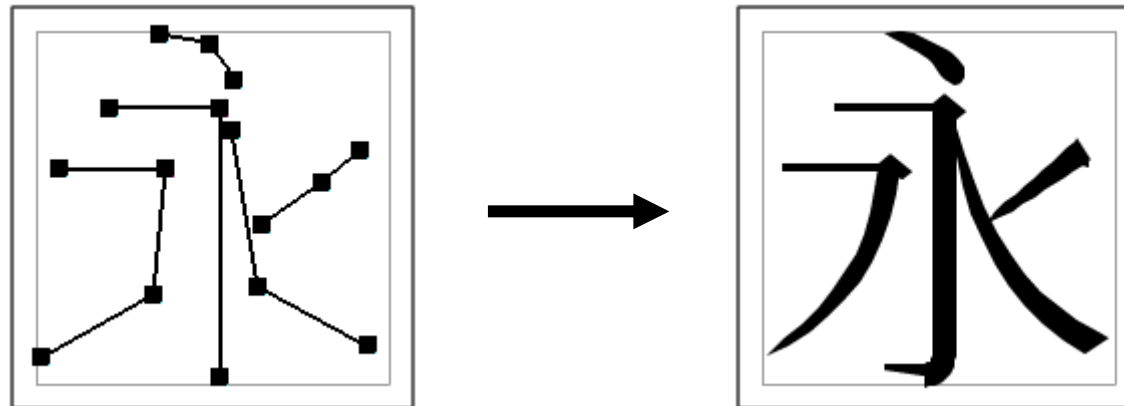
↑ shape of end of the stroke

↑ ↑ ↑
 coordinate X-axis, Y-axis (2 - 4 pairs)



Using KAGE system (4)

- modeling engine
 - generate *Song-ti* glyphs from skeleton data





Using KAGE system (5)

- use glyphs as components

- 99:0:0:11: 7 80:182: u8a00 (言)
- 99:0:0:63:10 186:190: u6c38 (永)

↑ use component

↑ coordinate of left top corner

↑ coordinate of right bottom corner

↑ name of component glyph





Using WikiWikiWeb

- wiki : web page which users can edit the document on browser just viewing at that time
 - anyone can join to edit
 - collaboration, divide of work
 - cf. Wikipedia : free online encyclopedia
- manage one glyph per one page
 - Wikipedia : 1 page 1 word
 - GlyphWiki : 1 page 1 glyph



Free license

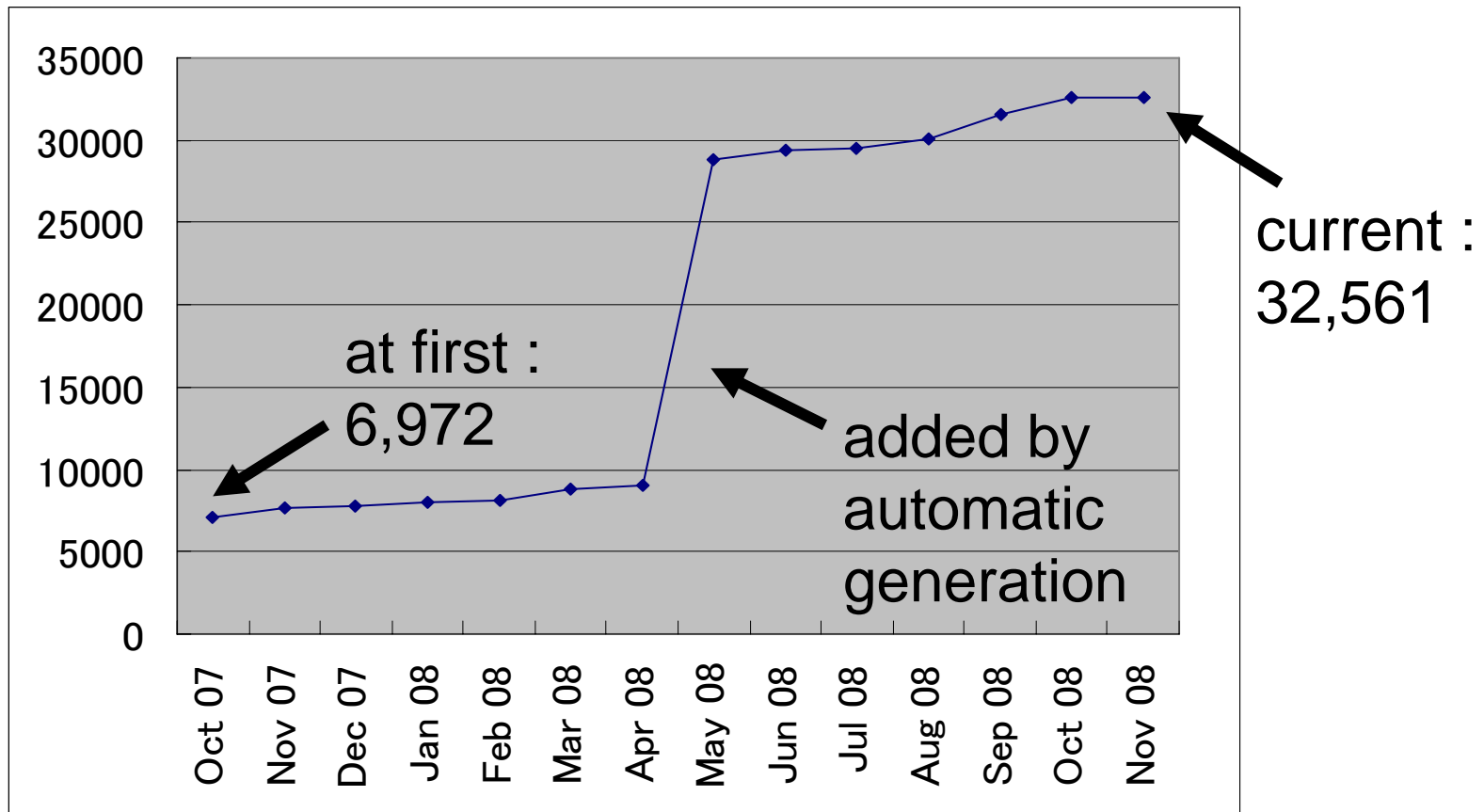
- GlyphWiki is free
 - (1) not necessary to pay : 免費
 - (2) no restrictions to use : 自由
 - copy, distribute with or without modification, either commercially and noncommercially
- you can create new glyph by processing registered glyphs as components

Sample glyphs registered in GlyphWiki

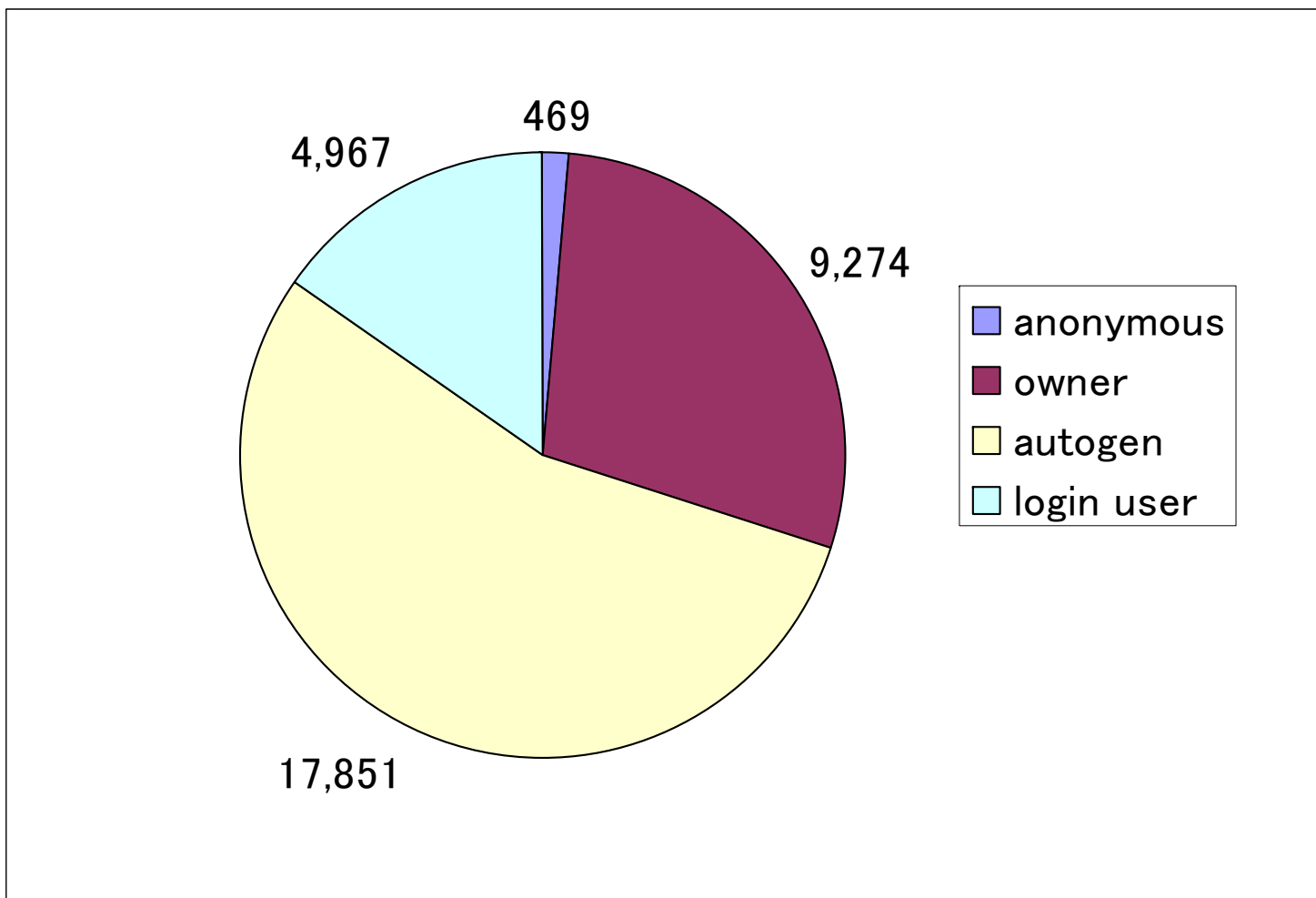


- ISO/IEC 10646
 - URO 一丁丂七丄丅 龠龡龢龣龤龥
 - add. 熻曙匪薪赴幹鋼驥驥鉞乚凶鏞堯マ
彡圭止出矣卓緜壻燠燠荏醫誦軻映
 - Ext. A 止丙崑今丰メ
 - Ext. B 丂丄丅丆万丈三上下
 - Ext. C 亂祺臻忙佻佻佻佻佻
 - Var. 骨骨 既既既既既
- Japanese Hanzi 氫体鬱樾腩底弄芑杏奥
- Joke 雨

A change of the number of glyphs



who registered glyphs



Releasing data as a font file



- June 2007 : 1st release
 - including 6,356 *Hanzi*
- October 2008 : 2nd release
 - including 10,204 *Hanzi*
 - distributed as Linux package (Debian/GNU Linux)

Demonstration





Future works

- continuing collection of glyphs
- internationalization of user interface
 - English (75%), Chinese, Hangul, etc.
- application to CCS3 (@font-face : embedding remote font to web pages)
 - users can see documents using PUA font on a browser without installing the font into own PC
 - Firefox 3.1, Safari 3

Thank you ! 多謝 !

