1   Proposal for a Japanese Script Root Zone LGR

2   *LGR Version 5*

3   *Date: 2021-9-30*

4   *Document version:J-LGR-0.20a*

5   *Authors: JGP (Japanese Generation Panel)*

6

7   Contents

22

23

24

25

26

27

28

1 General Information/ Overview/ Abstract

The purpose of this document is to give an overarching view of the label generation rules for the Japanese scripts including rationale behind the design decisions taken. This includes a discussion of the relevant features of the scripts, the communities and languages using it, as well as the process and methodology used and information of the contributors.

The formal specification of the LGR can be found in the accompanying XML document:

proposal-japanese-lgr-30sep21-en.xml

Labels for testing can be found in the accompanying text document:

japanese-test-labels-30sep21-en.txt

All the appendices to the document can be found in the accompanying EXCEL and PDF documents.

- Appendix A: Repertoire of J-LGR,
  Repertoire-LGR-Jpan-20210624.xlsx
- Appendix B: RESEARCH PAPER: SURVEY ON THE USER PERCEPTION OF THE HOMOGRAPHIC CHARACTER SET SPECIFIED BY JGP,
  ICANN-report-20200928.pdf
- Appendix C: Report of "Field Survey on visually identical character pairs",
  field-research-for-additional-identicalness.pdf

2   Script for which the LGR is proposed

ISO 15924 [a] Code: Jpan

ISO 15924 Key Number: 413

ISO 15924 English Name: Japanese (alias for Han + Hiragana + Katakana)

Latin transliteration of native script name: Hanzi, Kanji, Hanja, Hiragana, and Katakana

Native name of the script: 漢字(Kanji), ひらがな(Hiragana), カタカナ(Katakana),

Collectively called as 日本語 aka 和文 (Nihongo aka Wabun, Japanese) [b]

Maximal Starting Repertoire (MSR) version: MSR-5 [c]

To understand the basic background of above definition, it is worth noting that Japanese (Jpan) script is a composite script that covers three element scripts (Han (hereafter referred to as Kanji), Hiragana, and Katakana) as defined in ISO15924. In addition, it should also be noted that there is a baseline rule stipulating that IDN TLD labels must exclude alphanumeric characters and the hyphen, although some strings of Japanese words, including trademarks and trade names, contain alphanumeric characters in them.

---

[a] http://unicode.org/iso15924/iso15924-en.html
[b] https://en.wikipedia.org/wiki/Japanese_writing_system
[c] https://www.icann.org/sites/default/files/packages/lgr/msr/msr-5-wle-rules-06apr21-en.xml

3   Background on Script and Principal Languages Using It

(1) Background of scripts, characters, and languages

Japanese language is the only formal language of Japan. It is virtually the only native written language used in Japan, and is rarely used in other countries. Therefore, the population of writing system of Japanese language is about 125 million, which is the population of Japan as of January 2021.

Major scripts used for writing modern Japanese are Kanji, Hiragana, Katakana and alphanumeric (Latin alphabets and Arabic numerals). Characters in those scripts are used in a mixed way within one Japanese word, such as "A5 ランクの牛肉" (meaning "beef ranked as A5"), where 'A' is alphabet, '5' is numeric, 'ラ', 'ン', and 'ク' are Katakana, 'の' is Hiragana, and '牛' and '肉' are Kanji. Among such scripts, Kanji, Hiragana and Katakana are only allowed in Japanese TLD labels.

Kanji was imported from China in around the 5th Century AD. Hiragana and Katakana are phonetic syllabaries that were invented in Japan in around the 10th Century AD from cursive forms or parts of Kanji. Hiragana is mainly used as suffixes to Kanji to complete the full reading of the word, for adverbs, conjunctions, and to rewrite difficult Kanji into forms for easy writing and reading. Katakana is mainly used to represent loanwords and onomatopoeic words.

Modern characters in Kanji, Hiragana, and Katakana are defined in JIS (Japanese Industrial Standards) X0208 [d], which is certified by the government of Japan. For Hiragana and Katakana, they are only used in Japanese writing system and their repertoires are uniquely defined in JIS X0208. In JIS X0208 Kanji characters are categorized into two levels – frequently-used characters are at the first level (2,966 characters), and less frequently-used characters including those often used in personal names and geographic names are at the second level (3,390 characters). Many Kanji characters are shared with Chinese and Korean writing systems.

For about 360 pairs (or triplets or quadruples) of Kanji characters, there is a notion of an old form and a new form of the same Kanji character.[e] For example, '学' is a new form of '學 and they have the same pronunciation and meaning. In perceiving old form and new form of a

---

[d] https://en.wikipedia.org/wiki/JIS_X_0208
[e] https://en.wikipedia.org/wiki/Ky%C5%ABjitai and
    http://www2.japanriver.or.jp/search_kasen/search_help/refer_kanji.htm

121 Kanji character, Japanese people are divided into two types - those who think they are the
122 same character with different shapes, and those who think they are different characters and
123 each has its own independent status. Both are non-negligible in the population. In addition,
124 whether each individual person considers old/new form characters are the same or not varies
125 for each Kanji character. For example, depending on her/his circumstances, a person may
126 consider new form '学' and old form '學' are the same character but new form '応' and old
127 form '應' are different characters.

128

129 (2) Other relevant backgrounds

130

131 Words written in Kanji have phonetic mapping(s) to Hiragana and Katakana. To input a
132 Japanese Kanji string into computers, users usually type such Hiragana or Katakana string that
133 makes pronunciation of the intended string, then options for Kanji strings with such
134 pronunciation are displayed, and then the user selects Kanji strings (s)he wants to have in the
135 text. This means that multiple different Kanji words often have the same pronunciation. And
136 some Kanji strings have multiple pronunciations – i.e., have multiple mappings to Hiragana or
137 Katakana.

138

139

140

141

142

143

144

145

146

147

148

149

150

151

152

4   Overall Development Process and Methodology

4.1   Previous work

JP ccTLD registry started registration services of Japanese IDNs, called "Japanese JP domain names", in February 2001. In order to develop the rules for Japanese JP domain names, JPNIC [f], JP ccTLD Registry at the time, convened a task force populated from various experts such as domain name experts, trademark experts, character code experts and so on in September 2000, and developed rules for Japanese JP Domain Names. The rules are registered in IANA IDN Practices Repository as .jp Japanese (Japan) [g]. The rules are also adopted by IDN registration services in other TLDs such as .asia.

During the development process, specifications were published to the community and finalized through public comment process. Major characteristics of the specifications of Japanese JP domain names are defined as follows:

−   A domain label string consists of alphanumeric, Kanji, Hiragana and Katakana that contains one or more Kanji, Hiragana or Katakana.
−   Kanji must be in range of JIS X 0208 first level and second level (6355 characters). Moreover, Hiragana (85 characters), Katakana (89 characters) and quasi-Kanji (5 characters) must be in the same range.
−   No variants between characters exist.

As of 1 June 2021, 91 thousand (5.5%) of the 1.64 million JP domain names are Japanese JP domain names. During 20 years' experience of service delivery, there have been no complaints or objections to the Japanese Domain Name rules.

4.2   Diversity of Generation Panel Members

The members of the Japanese Generation Panel (JGP) consist of experts with various backgrounds. Each member has experience in one or more areas of IDN standardization, discussion between Chinese, Japanese, and Korean (a.k.a. CJK) IDN experts, establishment of Japanese Domain Name rules, registry business, registrar business, policy making in ICANN through participation from various sectors, and so on. For the list of all members, refer to Section 9.

---

[f] https://www.nic.ad.jp/en/
[g] https://www.iana.org/domains/idn-tables/tables/jp_ja-jp_1.2.html

187 ## 4.3   Work Process
188

189 Proposal of Japanese Root Zone LGR (hereafter referred to as J-LGR) was developed through
190 the following process shown in (1) - (7).

191 As CJK LGRs are inter-dependent in determining the repertoire, variants and WLE, frequent and
192 periodical discussion and coordination among CJK GPs (each referred to as CGP, JGP, and KGP)
193 and IP (Integration Panel) have helped CJK GPs a lot to maneuver to their final proposals.

194

195 (1) Establishment of JGP
196

197 Japanese Generation Panel (JGP) was informally formed and started its work in August 2014.
198 Early 2015, it submitted "Proposal for Generation Panel for Japanese Label Generation Rules for
199 the Root Zone" [h] to ICANN and formally acknowledged as one of the Generation Panels [i].

200

201 (2) Determination of initial repertoire and WLE as a starter

202

203 The process to design the current Japanese JP domain name rules and their usage were studied.
204 As the result of the study, it was concluded that the current Japanese JP domain name rules
205 were mostly adequate for TLDs. To be more precise:

206

207 As to the repertoire, JIS X 0208 is decided to be the repertoire of J-LGR following Japanese JP
208 domain name rules.

209

210 As to WLE, JP domain name rules allow any permutation of characters in the repertoire.
211 However, it is decided that more restricted rule be applied to the Root zone. Such rule is that
212 any iteration mark, prolonged mark, or small kana is prohibited to be the first character of
213 labels.

214

215

216

217

---

[h] https://www.icann.org/en/system/files/files/japanese-lgr-proposal-17mar15-en.pdf
[i] https://www.icann.org/news/announcement-2015-03-17-en

218   (3) Definition of the variants as a starter

219

220   The process to design the current Japanese JP domain name rules and their usage were studied.
221   As the result of the study, it was concluded that all the characters in the repertoire are
222   regarded independent. I.e., no variants were defined in J-LGR per se.

223   In the study, necessity of variant definition of old form and new form of a Kanji character was
224   revisited. And it was decided by JGP that there should be no variants. The reason was that
225   considering all characters as independent would be more appropriate because of its more
226   flexibility in choosing a TLD string to apply for. Additionally, those who want to use a set of
227   variant labels to designate one thing are not prohibited to apply for and use all of them. This
228   means that JGP decided the rule of "Japanese JP domain names should have no variants" still
229   stood.

230   At this stage, only the necessity of definition for variant characters with the same pronunciation
231   and meaning was considered. Visual identicalness, which will be discussed in (6), was out of
232   scope.

233

234   (4) Coordination between CGP, KGP and JGP in defining variants

235

236   From its initial phase, JGP has been consistent in its intention to accept the variants defined in
237   Chinese and Korean Root Zone LGRs in order to nurture safe TLD space.

238   This means that J-LGR imports all variants from Chinese LGR (hereafter referred to as C-LGR [1])
239   and Korean LGR (hereafter referred to as K-LGR [1]) when those three LGRs are merged into a
240   single Root Zone LGR. To assure that this import is adequate, JGP needed to observe and check
241   the process and result of the definition of variants made by CGP and KGP from time to time.

242   Adequateness was checked mainly from the following aspects.

243   (a) Not being too prohibitive to Japanese words
244   (b) Not generating too many variant strings that bring different meanings from one
245   Japanese word

246   Observing the process and result of CGP and KGP, which was discussed within JGP as well,
247   definition of C-LGR and K-LGR are considered to be adequate at some level. However, there was
248   an issue that needs consideration as discussed in (5).

249

250

251

252

253   (5) Reduction of the number of allocatable variant labels

254

255   Through (2)(3)(4) above, it has become known that the number of mutually-variant labels can
256   be big for some Japanese TLD labels because a lot of variant definitions would be imported by
257   merging CJK LGRs. As it is considered prohibitive that too many variant labels become
258   allocatable, JGP investigated the followings :

259   (A) necessity of making variant labels allocatable,   and
260   (B) measures to reduce the number of allocatable variant labels
261   (C) balance between "(A) necessity" and "(B) smallness of the reduced number"
262
263   (A) necessity of making variant labels allocatable

264   It was found that many old/new form relationships in Japanese Kanjis are the same as
265   traditional/simplified relationships in Chinese Han. Let's use '應' and '応' as an example.
266   Although original J-LGR defines old form '應' and new form '応' as independent characters, they
267   become a variant of each other when J-LGR and C-LGR are merged because '應' and '応' are
268   defined as variants in C-LGR. However, there is a possibility that this situation benefits Japanese
269   TLD applicants, especially those who want to respect old/new form relationship.

270   If an applicant considers '應' and '応' the same and wants to use both of them interchangeably,
271   (s)he may apply for a TLD containing '應' and also want to create a variant TLD containing '応'
272   at the same time. Actually under .jp, Keio University registers all old/new form combinations
273   慶応義塾大學.jp, 慶応義塾大学.jp, 慶應義塾大学.jp, and 慶應義塾大學.jp, and makes the
274   website of its university accessible from all types of old/new form lovers.

275   Making variant labels allocatable is desirable for Japanese TLDs to be usable and accessible.
276   Unless variant labels are allocatable, Keio University may give up application of its university
277   name as a Japanese TLD and have to apply for its university name as Chinese TLD if it wants to
278   use both '慶応義塾大学' and '慶應義塾大學'. Thus, it's desirable for J-LGR to avoid this
279   situation.

280   (B) measures to reduce the number of allocatable variant labels

281   Any permutation of characters in the repertoire are allowed in Japanese words. If this
282   characteristic is preserved in allowing variant labels to be allocatable, the number of allocatable
283   variant labels will go huge. To reduce the number of allocatable variant labels in a systematic
284   way, diverse mechanisms were devised and evaluated.

285   (i)   Making variant labels containing only variants that are Joyo-Kanji [j] (about 2,600
286   Kanji characters for everyday use) allocatable
287   (ii)   In addition to the above, making variant labels containing only 3 or less
288   characters that have Joyo-Kanji variants allocatable

---

[j] https://en.wikipedia.org/wiki/Joyo_kanji

289   (iii)   Only allowing the applied-for label to be valid and blocking all variant labels

290

291   (C) balance between "(A) necessity" and "(B) smallness of the reduced number"

292   Although above measure (i) reduces the number of allocatable variants, the number can be still
293   huge. For instance, if an applied–for label has 10 characters that have 3 Joyo-Kanji variants
294   each, 59 thousand (3 to the 10th power) variant labels become allocatable.

295   Above measure (ii) drastically reduces the number of allocatable variants. For any label, the
296   number will not exceed 27 (3 to the 3rd power). However, 27 is considered still big.

297   After every intensive effort, JGP came to a conclusion that "to make the number of allocatable
298   variant labels to 1 or 2" was just "to make static rules to select one or two label(s) that satisfy
299   any possible applicants regardless of their perception of old/new forms of characters" and it
300   was not implementable.

301   Thus, JGP decided that measure (iii) is the only and simple way to take for the Japanese
302   language community.

303   Although (ii) was deemed not to be adopted as J-LGR due to insufficient reduction of the
304   number, it is archived at JGP website [k] for future reference since this J-LGR version embracing
305   measure (ii) was the proposal that had accumulated every devise JGP had made so far.

306

307   (6) Reduction of confusion caused by visually identical characters

308

309   In Japanese JP domain names, all the permutations of characters in the repertoire are allowed
310   as domain labels. And the experience of Japanese JP domain names so far has observed no big
311   concerns related to visual identicalness of the labels. However, some concerns may be raised at
312   the TLD level.

313   It may reduce user confusion, if usage of visually identical labels is limited. To investigate if this
314   is practically true in case of Japanese scripts, field research was executed with human eyes, and
315   identically-looking characters were picked up. As a result, several pairs of characters were
316   decided to be deemed variants so that multiple visually identical labels do not co-exist as TLDs.

317

318

319

320

321

---

[k] https://j-gp.jp/J-LGR-v0.17a

322   (7) Create XML LGR for Japanese LGR proposal

323

324   The JGP creates the J-LGR in XML format following the RFC7940 [2].

325   The XML LGR was tested and verified using ICANN LGR Tool [I] and test labels. (japanese-test-
326   labels-30sep21-en.txt)

327

[I] https://lgrtool.icann.org

328 # 5   Repertoire

329

330 The Repertoire of J-LGR is shown in Appendix A. (Repertoire-LGR-Jpan-20210624.xlsx)

331

332 As the usual basic set of Japanese characters is defined by JIS X 0208 and it has a successful

333 track record under Japanese JP Domain Name experience, it is decided to have JIS X 0208 as the

334 repertoire of J-LGR.

335 The repertoire of Japanese LGR (J-LGR) is defined as below.

336   (1) Kanji (Han/Hanja)

337   The first level and the second level Kanji defined in JIS (Japanese Industrial Standard) X

338   0208:1997 (6356 characters in the range of U+4E00-U+9FA0, plus 2 characters in the

339   range of U+3005-U+3006, 6358 in total). All of them are included in MSR-5.

340   (2) Hiragana

341   The first level Hiragana defined in JIS X 0208:1997 (83 characters in the range of

342   U+3041-U+3093, plus 2 characters in the range of U+309D-U+309E, 85 characters in

343   total). All of them are included in MSR-5.

344   (3) Katakana

345   The first level Katakana defined in JIS X 0208:1997 (86 characters in the range of

346   U+30A1-U+30F6, 3 characters in the range of U+30FC-U+30FE, 89 characters in total). All

347   of them are included in MSR-5.

348

349 As a result, J-LGR contains 6532 Japanese characters in its repertoire (hereafter referred to as

350 original repertoire).

351

352 C-LGR and K-LGR define variants sets which include Japanese characters described in (1)-(3).

353

6 Variants

As all the characters are generally regarded independent and the experience of Japanese JP domain name services has observed a successful track record, it was decided that no variants, except a small number of visually identical characters described below, are defined in J-LGR at the beginning. However, after variants are defined in C-LGR and in K-LGR, J-LGR imports all those variant definitions, so that Chinese or Korean language community won't be confused by rejecting their required variant definitions. Hereafter referred to as Origin-1.

From a different aspect, there is an issue where visually identical characters should be handled adequately to avoid user confusion. It's typically solved by making mutually visually identical characters 'variants'. Consideration given and the resulted definition of variants are shown in Section 7. Hereafter referred to as Origin-2.

As a consequence, an applied-for Japanese label may generate multiple variant labels, the number of which may logically exceed tens of thousands. This multiplicity occurs mainly due to import of variant definitions from Chinese LGR and Korean LGR, although every character in Japanese original repertoire is treated as mutually independent. All imported and visually identical variants are blocked.

The following table shows the origin of blocked variants used in J-LGR.

Type	Origin	Comment
blocked	1	Imported variant character
blocked	2	Visually Identical variant character

(note) The distinction between "imported variant character" and "visually identical variant character" is documented using ref="101" for visually identical variant characters or ref="300" for imported variant characters in XML file.

7 Visually Identical Characters

Consultation with Root Zone LGR panel community, especially based on discussion between IP and CJK GPs, JGP decided to incorporate in-script visually identical characters as variants if there were objective evidence.

There are two kinds of cases for visual confusability as follows:

(1) One stroke mark character and Kanji

(2) In-Japanese-scripts (between Hiragana, Katakana and Kanji) visually identical characters

Each is explained below.

(1) One stroke mark character and Kanji

Following two pairs are defined to be variants because they resemble each other, they are single stroke, and one of them is a mark character.

Code Point	Glyph	Name	Code Point	Glyph	Name
U+30FC	ー	Katakana-Hiragana Prolonged Sound Mark	U+4E00	一	CJK Unified Ideograph
U+30FD	ヽ	Katakana Iteration Mark	U+4E36	丶	CJK Unified Ideograph

(2) In-Japanese-scripts (between Hiragana, Katakana and Kanji) identical characters

(a) Initial set of visually identical character pairs

JGP selected eight candidates from "confusable mapping for IDN" recommended by UNICODE Consortium [m]. The reason why JGP used confusables.txt is that it was universally well defined as a deliverable of the Unicode Consortium's work. Selection of visually identical pairs of characters in J-LGR repertoire is made with the following criteria using confusables.txt:

---

[m] https://www.unicode.org/Public/security/latest/confusables.txt

406 − Code point at the most left column is inside JGP's original repertoire (hereafter referred to
407   as code1)
408 − One or more code point(s) of the second left column of code1 is inside JGP's original
409   repertoire (hereafter referred to as code2)
410 − At least, one of code1 and/or code2 is Hiragana or Katakana

411 Using the above criteria, eight pairs are extracted.

412 With those eight pairs and some additional conditions such as various font set, font size and
413 examinees' language background, JGP asked ICANN to conduct a field research by a third party.
414 JGP received research report shown in Appendix B which concluded that it's difficult to
415 distinguish all eight pairs in any font, font size or language background of the subjects. As a
416 result, JGP decided to define those eight pairs as variant based on visual identicalness. Also JGP
417 decided to add two more pairs – i.e., visually identical characters with visually identical marks –
418 which should be obviously relevant. As a result, JGP defined ten pairs as variants shown below.

419

420 Following ten pairs are defined to be variants because they resemble each other.

Code Point	Glyph	Name	Code Point	Glyph	Name
U+3078	へ	Hiragana Letter He	U+30D8	ヘ	Katakana Letter He
U+3079	べ	Hiragana Letter Be	U+30D9	ベ	Katakana Letter Be
U+307A	ぺ	Hiragana Letter Pe	U+30DA	ペ	Katakana Letter Pe
U+30A8	エ	Katakana Letter E	U+5DE5	工	CJK Unified Ideograph
U+30AB	カ	Katakana Letter Ka	U+529B	力	CJK Unified Ideograph
U+30BF	タ	Katakana Letter Ta	U+5915	夕	CJK Unified Ideograph
U+30C8	ト	Katakana Letter To	U+535C	卜	CJK Unified Ideograph
U+30CF	ハ	Katakana Letter Ha	U+516B	八	CJK Unified Ideograph
U+30CB	ニ	Katakana Letter Ni	U+4E8C	二	CJK Unified Ideograph
U+30ED	ロ	Katakana Letter Ro	U+53E3	口	CJK Unified Ideograph

421

422   (note) "U+3079 and U+30D9" and "U+307A and U+30DA" pairs are the added variants derived
423 from "U+3078 and U+30D8" pair.

424

425

426   (b) additional definition of visually identical character pairs

427   Further investigation was conducted on the necessity of adding more pairs (or sets) to those in (a) as
428   visually identical pairs in J-LGR. As measures for this, a field survey was conducted to see whether the
429   Internet users misunderstood characters due to visual identicalness of pairs other than those in (a). The
430   report of this research is in Appendix C.

431   According to the result of the survey, some survey respondents said some pairs were confusing enough
432   to be misunderstood because of their visual identicalness. However, for every such pair, the percentage
433   of those respondents who made misunderstanding was less than 3% of all the respondents. Therefore, it
434   is concluded that it is appropriate for J-LGR to define only those 10 pairs in (a) as visually identical and to
435   be treated as variants.

436

8   Designing Dispositions and Whole Label Evaluation Rules (WLE)

As stated in Section 4.3(5) and Section 6, after a long and exhaustive discussion in Root LGR community, JGP decided to reduce the number of allocatable labels by blocking any variant labels.

In detail, allocatable variant labels are determined as below:
- Original label (a label that is applied-for itself) is valid
- Other variant labels containing at least one variant character that is visually identical or imported Kanji variant are blocked

JGP defines 1 special rule for WLE.

Defined rule #1

   Any small kana, iteration mark or prolonged mark must not start a label.

453 # 9 Contributors

454

455 Notable contributors in developing J-LGR

456 JGP members

457   Hiro Hotta (Chair)   JPRS

458   Akinori Maemura (Vice Chair) JPNIC

459   Shigeki Goto   Waseda University and JPNIC

460   Kazunori Konishi   APAN

461   Tsugizo Kubo   Senshu University

462   Yoshitaka Murakami   Com Laude Japan

463   Shuichi Tashiro   Information-technology Promotion Agency, Japan

464   Yoshiro Yoneya   JPRS

465   Yuri Takamatsu (Secretary) JPRS

466

467 CGP Members (among others, only Chairs are listed)

468   Kenny Huang (co-Chair)

469   Wei Wang (co-Chair)

470

471 KGP Members (among others, only Chair is listed)

472   Kyongsok Kim (Chair)

473

474 IP members

475   Asmus Freytag

476   Marc Blanchet

477   Michel Suignard

478   Nicholas Ostler

479   Wil Tan

480   ICANN staff

481   Sarmad Hussain

482   Pitinan Kooarmornpatana

483   Jianchuan Zhang

484

485

486   # 10 References
487

[1] Proposals for Root Zone Label Generation Ruleset,
    https://www.icann.org/resources/pages/lgr-proposals-2015-12-01-en

[2] K. Davies, A. Freytag, "Representing Label Generation Rulesets Using XML", RFC 7940,
    https://www.rfc-editor.org/rfc/rfc7940.txt

488

489

490 ## Appendix
491

492 **Appendix A: Repertoire of J-LGR,**
493 **Repertoire-LGR-Jpan-20210624.xlsx**
494 The EXCEL document includes 6532 JGP Unicode code points.
495

496 **Appendix B: RESEARCH PAPER: SURVEY ON THE USER PERCEPTION OF THE HOMOGRAPHIC**
497 **CHARACTER SET SPECIFIED BY JGP,**
498 **ICANN-report-20200928.pdf**
499 The PDF document is research report of survey regarding user perception of visually identical
500 characters in Japanese scripts. This survey was conducted by ICANN and Waseda University.
501

502 **Appendix C: Report of "Field Survey on visually identical pairs",**
503 **field-research-for-additional-identicalness.pdf**
504 The PDF document is research report of survey regarding Internet users' experience of
505 misunderstanding due to visual identicalness of characters in Japanese scripts. This survey was
506 conducted by JGP.