

# 1        Consideration on reduction of allocatable labels (v0.1)

## 2 3    Background

4  
5        CGP, JGP, and KGP are collaborating to create Root zone LGR for CJK. CJK GPs'  
6        collaboration is not limited to among us, but also involving IP and ICANN. Because  
7        Han characters have several variants in certain language(s), the integrated LGR  
8        (LGR-2, c.f. terminology) may generate tens of variant labels. Inflation of Root zone  
9        by variant labels are not desired, therefore some mechanisms for reducing allocatable  
10       labels must be considered.

## 11 12   Problem statement

13  
14       Many of Han characters in LGR-2 have 2 or more variants. If an applied-for label  
15       has 5 characters and each character has 2 variants, it generates 32 ( $=2^5$ ) variant  
16       labels including applied-for label<sup>1</sup>. According to the proposed CJK integration  
17       algorithm, all the variants defined by other GPs are imported and make their variant  
18       type to “allocatable” except for the cases where it is not defined in original repertoire.  
19       Therefore, for example, if all of them are allocatable and applicant requests and is  
20       allowed to delegate most of them, the Root zone size easily inflates. This situation is  
21       not desired, so reduction mechanism of allocatable variant labels in Root zone LGR is  
22       highly recommended. In addition, labels in Root zone should be predictable from end  
23       users' perspective. This requires that set of variant labels in Root zone must be  
24       consist from combination of variants used in daily life, and keep small as much as  
25       possible. The “allocatable” reduction mechanism also must take this into account.

26  
27       Followings are some possible reduction mechanisms:

- 28  
29       (1) ICANN/IP prepares additional process for applicant or evaluation-after-  
30       application panel to select reasonable numbers of variant labels from generated  
31       “allocatable” variant labels.  
32       (2) Each GP struggles to reduce variants in LGR-1, for instance by reducing variant  
33       characters.  
34       (3) After the first CJK integration process (first generation), each GP reproduce next

---

<sup>1</sup> It is unclear whether “applied-for label” is always allocatable.

generation LGR-1 with passive variant definitions imported from first generation  
LGR-2 and some proper variant subtypes for them.

## Considerations for each possible mechanism

For (1), this mechanism requires human interaction which takes extra time and resources. Root zone LGR is intended to provide automatic method as much as possible, so there will be a contradiction. We will need further discussion to this mechanism. Note that this human interaction process will keep “delegated” variant labels in predictable range.

For (2), CGP had already employed their best effort to reduce variants. JGP defined no variants. KGP are going to define very small set of variants. Therefore, it seems unfeasible or unrealistic to do this.

For (3), CGP had already defined simp/trad/both variant subtypes and prohibits mixture of simp/trad subtypes. JGP investigated probability of numbers of generated variant labels and actual usage of variant labels from existing “Japanese.JP” domain names. And JGP tentatively concluded that definition of proper variant subtype is impossible as there seems to be no general common and predictable rule for applied-for-string regarding variants<sup>2</sup>. We will need KGP’s feedback.

---

<sup>2</sup> Statistics will be provided during the meeting.