

## Latin Generation Panel (GP) Meeting

### Notes from the meeting on 2 July 2020

Meeting Attendees (in alphabetical order)

GP members:

1. Bill Jouris
2. Dennis Tan
3. Hazem Hezzah
4. Mats Dufberg
5. Meikal Mumin
6. Michael Bouland
7. Mirjana Tasic

Latin script users

1. Cengiz Acarturk

Staff:

2. Pitinan Kooarmornpatana
3. Sarmad Hussain

### Meeting Notes

The GP used the Variant Discussion for I and dotless I presentation for discussion. Cengiz presented an example for Turkish users as follow:

An IT company 'issız communications' wants to apply for a TLD 'issız', while a restaurant 'issiz restaurant' wants to apply for a TLD 'issiz'.

- Both "issız" and "issiz" are legitimate words in Turkish. They have different meanings.
  - The communications company (issız) would be happy to have "issiz" as TLD, too, due to current habit of using "i" instead of "ı" in Turkey.
  - The restaurant will not be interested in having "issız" as TLD, in addition to "issiz", since nobody would use "ı" instead of "i" in a domain name.

Cengiz shared that "i" and "ı" should be variants. The Turkish users are used to the habit of using "i" instead of "ı", therefore the label using "i" is ambiguous with the same label using "ı".

Based on the example, one of the the possible solutions was to make them variants with mapping from dotless I → I allocatable, and mapping from I → dotless I blocked.

It was noted that the "i" and "ı" case is unique as it loses the dot when lowercasing while other characters with dot in Turkish maintain the dot when they are lowercased.

It was raised that in Swedish passport, the name part can be non-ASCII but its transliterated ASCII characters are used in the machine readable area. Cengiz shared that it is different for Turkish, when applying for VISA, the applicant will automatically use the ASCII characters for their names, "i" is used instead of "ı".

The GP raised that in browsers the lowercasing is consistent independent on locale. The capital “ı” will be lowercased to “i with another dot”. And the capital İ will be lowercased to “i”. Cengiz shared that if use uses capital letter, they will naturally use capital capital İ to reach the correct address.

A GP member mentioned that the motivation of defining them variants needs to be discussed further. The linguistic rationale should not be used. Linguistic rationale would also be applicable to many other cases in Latin script.

Another GP member noted that as there are strong cases of confusion for the local users, it should be made variants.

Cengiz shared another example; when Turkish users buy something online and enter the name and surname for credit card information, the form always accepts “ı” as a variant of “i”. These two letters in the offline world are different, but in the IT system, these are variants in various platforms e.g. banking. The Turkish users will be surprised if a Turkish webform treats these letter differently.

It was raised that when the GP discussed sharp S and ss, the GP only analyze from the perspective of German context as the Sharp S is only used in Germany, making Sharp S and ss variant does not affect users outside Germany. This might be a similar case. The “ı” is only used in Turkish and Azerbaijan, therefore it should be discussed from the Turkish and Azerbaijan context.

It was reiterated that the down-casing behavior increase the risk of confusion. Though the browsers are consistently manage the two characters but user could end up in the different website from the intended one.

Dennis volunteered to research more on the topic to find the rationale which is not linguistic based to avoid open up other concluded cases.

A GP member commented that end-user’s perception should be taken as priority. Another member raised that there was no cost to make them variants, as long as the number of generated variants is controlled.

The GP will continue discussion in the next meeting.

The GP thanked Cengiz and Kadir for attending the meetings and sharing his knowledge, examples, and other useful information.

**Next meeting:** Thursday 9 June 2020 16:00UTC.

#### **Action Items**

<b>S. No.</b>	<b>Action Items</b>	<b>Owner</b>
1	<i>Research further to find the variant supporting rationale which is not linguistic based to avoid open up other concluded cases.</i>	DT