

Concerns on Old Hanzi Activities

Dr. Atsushi Suzuki

Japan NB

Introduction

- Old Hanzi Activities
 - Standardization of JiaGuWen (甲骨文) Characters
- Characteristics of JiaGuWen
 - Used for 300 years (13th century B.C. to 11th century B.C.)
 - Some characters change greatly during this period.
 - Many glyphs are excavated, and nearly 6,000 characters are identified.
 - Many variants. (e.g. mirror images, ages, writer habits.)

Requirements for Encoding Old Characters

- Clarification for Target Users
 - Who will use it?
 - Archaeologist, Historian, Linguists, etc..
 - Perspectives of “characters” differs among people.
- Policy on Glyph Selection
 - Only “established” glyphs should be encoded.
 - Questionable glyphs to be filtered out.
- Establishment of Unification Criteria
 - Which glyphs should be unified to be single “character”?
- Character Arrangement and Encoding Schemes
 - How characters are arranged, and how variants are treated. (Different codepoints vs. Variation Selectors, etc.)

Current Old Hanzi Activities

- Target User is not specified.
 - This is crucial issue on establishing “unification criteria”.
- Glyph Selection Policy changes during standardization activities
 - Only one representative glyphs for each character are chosen at initial stages, but variants are actively adopted at later stages.
 - Reasons for glyph selections and character assignments are inconsistent
- Unification Criteria is not established either.
 - Character \Leftrightarrow Glyph abstraction policy is not solid.
 - No consideration for how to handle variants.
 - Which variants should be differently encoded, and which ones should be unified, is determined case by case.

Conclusion:

Proposal for Old Hanzi Activities

- **Specific Target Users, Policy on Glyph Selection, Unification Criteria, and Variant Encoding Scheme** must be clearly established, and thoroughly enforced during the entire standardization activity.
 - Experts on character encoding should be added to Old Hanzi Group.
- Collected JiaGuWen glyphs must be disposed in accordance with these policies.
 - Proper glyph selection and character abstraction should be performed.