Shape Data Sets

UNIVERSITY
OF COLOGNE

IDH INSTITUT FÜR
DIGITAL HUMANITIES
UNIVERSITÄT ZU KÖLN

Vyshantha Simha yshant@gmail.com

Research Studies - Work In Progress

Shape Data Sets

This document details of the various glyph data sets being attempted to be included in full (or part) based on authenticity and evidences.

It also details what appears to be the boundary of what should & shouldn't be included and what is intentionally excluded.

Conscripts including <u>Neography</u> created for an existing or extinct or new language (or for a particular niche), if & only if it is agreed by committee or group or panel to be used regularly then it would be included.
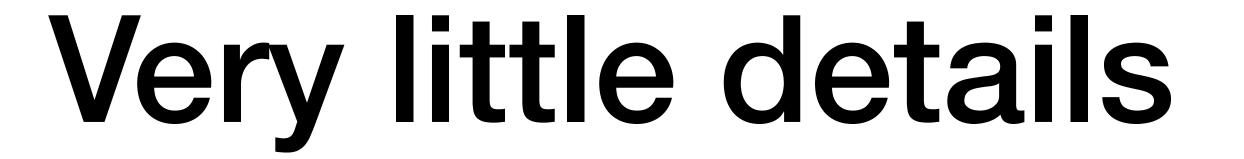
Boundary of inclusion

Historic scripts, fonts (including calligraphy), notations, ciphers, numerals and undeciphered/uncategorised sets

Fonts from digital age (after 1960s CE) for existing scripts are excluded and only font as face value for Unicode code points used & others for which they are available will be used

Very little details

- Archaic Miao Square Script
- Baima / Namuyi Tibetan hieroglyphs
- Cacaxtla
- Duota
- Izapa
- Jing (Zinan)

- Kabiye
- Kadu
- Komo Signs
- Sinuwat
- Tajin
- Takalik Abaj
- Tula

Yet to decide

- German dialectology (Teuthonista) , Transitus Fluvii , Toki Pona (Sitelen Pona, Sitelen sitelen) , Aksara tambo ruweh , Gar Dusun

- Hmong script (Pa Chai, Ntawv Savina, Romanised alphabets sets, Hmong Lao, Whitelock systems, Trung script, Miao Chinese characters)

- Cuban Abakuá tradition (Anaforuana or Gandó)

- Vanautu Sand-art, Vévé Haitian Vodou, Rangoli

- Fon-applique, Korhogo , Bogolanfini , Hajichi , Kodo-patterns , Tai Daeng Shaman, Hu-Ronko , Cenda , Dinagat , Goguryeo, Hmong Embroidery

- Aboriginal memorial (Funeral poles or *dupun*), Totempoles (similarly - Chemamüll, Serge, Tiki, Jangseung, Huabiao, Stogastulpis, Irminsul, Maypole), Lukasa memory board

- Tamga (clan symbols), deity symbols, Heraldry-quartering (Etiquette), Palmette , Wuyue Pictogram , Mayan emblems signs (other Emblems), Monogram (Roman coin) , Gelimu (Gemu, Daba, Gopa, Marimasa, Larko)

- Klingon (piqad) : large learning material & community support

Unified sets exclusion

Given Unicode sets contain mostly graphemes per script & are unified thus excluded independent sets of glyphs per variety - essentially independent **writing systems** aren't integrated. For examples :

Batak

Old Italic

Canadian Syllabic

…

Sets Excluded

- Scripts yet to be approved by committee

  <u>Sankethi (2021)</u>

  <u>Badagu (2009)</u>

  Bassa Syllabary (1836) : African Writing Systems of the Modern Age, 1st. Edition, 2011, A. Rovenchak, J. Glavy, p79

  Cherokee Ideograms

Sets Excluded

- Archeological findings of only 1 inscription :
  - Naxian Sphinx
  - Alekanovo inscription
  - Para-Lydian in Sardis
  - Tell Umm el-Marra (possible earliest alphabets)

Sets Excluded

- Archeological findings of only 1 document publication :
  Vellara
  Dialectal Palaeogtype (survey)
  MS 73525 (British Library)
  Lingua ignota
  Codex Seraphinianus
  Voynich (artificial language)
  Hamptonese
  Rohonc
  Dorabella

Sets Excluded

- Multi-font (or script) Dictionary & universal Phonology books by individuals :

  Avramović
  Franklin Phonetic
  Global Alphabet
  Srpski
  Zhelekhivka
  Universall Alphabet
  Lodwick
  Vtopiensium Alphabetum

Sets Excluded

• Items created for propaganda, or hoaxes, or to generate beliefs among followers
  Ahnenerbe (Ariosophy)
  Latter Day Saint movement artefacts :
    Kinderhook plates (Illinois, USA)
    Voree plates (Wisconsin, USA)
    book of Mormons archive like Elephant Slabs, Metcalf stone, Peruvian Gold
plate, Ecuador Copper plate, tablet of Parahyba, Tlatico cylinder seal Mexico
  pseudo-Phoenician from Glozel France
  Jindai moji or Kamiyo moji (Hokkaido, Woshite, Izumo, Ahriru, Ahiru kusa,
Tsukushi, Katakamuna, Ryukyu, Toyokuni, Tsushima)
  Alifuru

Sets Excluded

• Script "Pareidolia" or pará ($\pi\alpha\rho\acute{\alpha}$) -script ?

• Game signs (<u>Mahjong</u>, <u>Chess</u> , <u>Playing Cards</u>, <u>Domino</u>, etc…)

• Company logos and / or Trademark and Registrated icons

• Technology based : <u>Barcode</u> & QRcodes

• Biosemiotics (animal marks , fingerprints, palmistry)

• visual representation systems or graphical communication tools like maps as a whole (including <u>Hunting Kites</u>), diagrams, networks, charts, schematics, etc.

Sets Excluded

- For certain constructed/artificial <u>Languages</u> and some Script (<u>Omniglot</u> - <u>Skyknowledge Conscript</u> & <u>ConScript Unicode Registry</u>) created by individuals and not used in any context outside of said individuals or groups (like fictional & artistic writing : books, movie, plays, experiments, etc.)

    Aiha, Cirth, Sarati, Tengwar

- "Alien Writing" (seance communications), Sigil (angel or magic communications), Simic & Asemic writing

    <u>Magi</u> , <u>Angelic (Celestial)</u> , <u>Malachim</u> , <u>Enochian</u> (<u>Kelley, Dee, Crowley</u>) , <u>Écritures Martiennes</u> (<u>Martian</u>) , book <u>Pantographia</u>