മലയാളം കമ്പ്യൂട്ടിങ്ങ് : യുണിക്കോഡ് പ്രശ്നങ്ങളും സാങ്കേതികവിദ്യയും

സന്തോഷ് തോട്ടിങ്ങൽ, സ്വതന്ത്ര മലയാളം കമ്പ്യൂട്ടിങ്ങ്

യുണിക്കോഡ്

മലയാളം കമ്പ്യൂട്ടറുകളിൽ വ്യാപകമായി ഉപയോഗിക്കാൻ തുടങ്ങിയെങ്കിലും അടിസ്ഥാനപരമായ പല പ്രശ്നങ്ങളും ഇനിയും പരിഹരിക്കാനുണ്ട്. ഇവ പരിഹരിച്ചല്ലാതെ നമുക്ക് വളരെ മുന്നോട്ടു പോകാനാകില്ല. മലയാളം കമ്പ്യൂട്ടറുകളിൽ ഉപയോഗിക്കണമെങ്കിൽ വ്യക്തമായി നിർവചിക്കപ്പെട്ട ഒരു ഡാറ്റാ എൻകോഡിങ്ങ് സ്റ്റാൻഡേഡ് അത്യന്താപേക്ഷിതമാണ്. അക്ഷരങ്ങൾ ഒരോന്നും എങ്ങനെ വ്യതിരിക്തമായി രേഖപ്പെടുത്താം എന്നതിനുള്ള മാർഗ്ഗരേഖ. യൂണിക്കോഡ് മാനകമാണ് ഇതിനായി നിലവിലുള്ളത്. ഓരോ മലയാളാക്ഷരങ്ങൾക്കും യുണിക്കോഡിൽ അതിന്റെയായ കോഡ് പോയിന്റുകൾ ഉണ്ട്. ഇതുപയോഗിച്ചുതന്നെയാണ് നാം മലയാളം കമ്പ്യൂട്ടറുകളിൽ ഉപയോഗിക്കുന്നതും ഉപയോഗിക്കാൻ പോകുന്നതും. പക്ഷേ മലയാളം യൂനിക്കോഡിൽ ഗുരുതരമായ പ്രശ്നങ്ങളുണ്ട്. അധികം ആഴങ്ങളിലേക്കു പോകാതെ അവയെ താഴെ കൊടുത്തിരിക്കുന്ന വിധം ലളിതമായി അക്കമിട്ടു നിരത്താം.

1. കേരള സർവ്വകലാശാല, രചന അക്ഷരവേദി, സ്വതന്ത്ര മലയാളം കമ്പ്യൂട്ടിങ്ങ് തുടങ്ങിയവയിലെ ഭാഷാ വിദഗ്ധരും സാങ്കേതികവിദഗ്ധരും ശക്തമായെതിർത്തിട്ടും യുണിക്കോഡ് കൺസോർഷ്യം യൂണിക്കോഡ് 5.1 പതിപ്പിൽ കൊണ്ടുവന്ന മലയാള ചില്ലക്ഷരങ്ങളുടെ രണ്ടാം എൻകോഡിങ്ങ്. രണ്ടുവിധത്തിൽ ചില്ലക്ഷരങ്ങളെഴുതാമെന്നത് മലയാളം കമ്പ്യൂട്ടിങ്ങിന്റെ അടിത്തറയെത്തന്നെ ബാധിക്കുന്ന ഗുരുതരമായ പ്രശ്നമാണ്. സുരക്ഷാപ്രശ്നങ്ങൾ, തെരച്ചിൽ/അകാരാദിക്രമം എന്നിവയിലുണ്ടാക്കുന്ന പ്രശ്നങ്ങൾ, ഭാഷാ വ്യാകരണവ്യവസ്ഥക്ക് നിരക്കാത്തതിന്റെ പ്രശ്നങ്ങൾ തുടങ്ങിയവ ഇതുണ്ടാക്കുന്നു.
2. ന്റ എന്ന കൂട്ടക്ഷരം ൻ + ചന്ദ്രക്കല + റ എന്നെഴുതണമെന്ന യൂണിക്കോഡ് നിർദ്ദേശം. മൈക്രോസോഫ്റ്റിന്റെ കാർത്തിക എന്ന ഫോണ്ടിനുവേണ്ടി കൊണ്ടുവന്ന ഈ നിർദ്ദേശം ഭാഷ നിയമങ്ങളെ കാറ്റിൽ പറത്തുന്നു.
3. വർത്സ്യാനുനാസികവും(പനയിലെ ന), വർത്സ്യഖരവും(റ്റയുടെ പാതി, അല്ലെങ്കിൽ ഇംഗ്ലീഷിലെ cat എന്നു പറയുന്നതിലെ അവസാനത്തെ റ്റയുടെ രൂപം) യുണിക്കോഡിൽ യാതൊരു ചർച്ചയും കൂടാതെ തെറ്റായ തെളിവുകളുടെ പിൻബലത്തിൽ കൊണ്ടുവന്നത്.
4. രേഫത്തിനു പ്രത്യേകം കോഡ് പോയിന്റ് കൊടുത്തതിലൂടെ വന്ന സുരക്ഷാപ്രശ്നങ്ങളും സാങ്കേതികപ്രശ്നങ്ങളും. ഭാര്യ, സൂര്യൻ എന്നിവയൊക്കെ പല രീതിയിൽ എഴുതാൻ ഇതു കാരണമാവുന്നു.
5. ഔ ചിഹ്നത്തിന് ൌ, ൗ എന്നിവയിലേതുപയോഗിക്കണമെന്ന ആശയക്കുഴപ്പം.

മേൽപ്പറഞ്ഞ പ്രശ്നങ്ങൾ ഇതുവരെ ഉന്നയിച്ചതും അവയ്ക്കെതിരെ നിലകൊണ്ടതും സ്വതന്ത്ര സോഫ്റ്റ്‌വെയർ പ്രവർത്തകർ മാത്രമാണ്. കുത്തക കമ്പനികളുടെ പ്രതിനിധികൾക്കു മാത്രം തീരുമാനമെടുക്കാവുന്ന യുണിക്കോഡ് കൺസോർഷ്യത്തിൽ നമുക്ക് പ്രാതിനിധ്യമില്ല. അവരുടെ തീരുമാനങ്ങൾ സുതാര്യവും അല്ല. ഈ പ്രശ്നത്തെക്കുറിച്ച് ഭാഷാ വിദഗ്ധരെ അവബോധമുള്ളവരാക്കുകയും അവരുടെ അഭിപ്രായം യുണിക്കോഡിൽ ഉന്നയിച്ച് എത്രയും പെട്ടെന്ന് പരിഹാരമുണ്ടാക്കുകയും ചെയ്തില്ലെങ്കിൽ മലയാളം കമ്പ്യൂട്ടിങ്ങിന് അധികദൂരം മുന്നോട്ടു പോകാനാകില്ല.

ലിപിവ്യവസ്ഥ

അക്ഷരങ്ങളുടെ തലത്തിൽ നിന്നും വാക്കുകളുടെ തലത്തിലേക്കു വരുമ്പോൾ നിയതമായ എഴുത്തുരീതിയില്ലാത്തത് മലയാളത്തെ വലയ്ക്കുന്നു. പലവാക്കുകൾ പലരീതിയിൽ എഴുതാമെന്നത് വലിയൊരു വെല്ലുവിളിയാണ്. അദ്ധ്യാപകനും, അധ്യാപകനും കമ്പ്യൂട്ടറിനെ സംബന്ധിച്ചിടത്തോളം രണ്ടുവാക്കായിരിക്കുമ്പോൾ ഈ ഡാറ്റയുടെ മുകളിലെ സകല കണക്കുകൂട്ടലുകളും തെറ്റുന്നു. ഇതു പരിഹരിക്കാൻ ഏകീകൃതമായ ഒരു എഴുത്തുരീതി

അടിച്ചേൽപിക്കൽ പ്രായോഗികമല്ല താനും. ലിപിപരിഷ്കരണങ്ങളിൽ നിന്നും ഉടലെടുത്ത ഈ അവ്യവസ്ഥകൾക്ക് എങ്ങനെ പരിഹാരമുണ്ടാക്കാമെന്ന് ആലോചിക്കേണ്ടിയിരിക്കുന്നു.

സ്റ്റാൻഡേഡുകൾ

മലയാളഭാഷാ കമ്പ്യൂട്ടിങ്ങിന്റെ സ്റ്റാൻഡേഡുകൾ തീരുമാനിക്കുന്നതിൽ ഇതുവരെ കണ്ടുവരുന്നത് ഒട്ടും സുതാര്യമല്ലാതെ ജനകീയ പങ്കാളിത്തമില്ലാതെ ഏതെങ്കിലും സർക്കാർ ഏജൻസികൾ(ഉദാഹരണം:സിഡാക്ക്) യാതൊരു ഉത്തരവാദിത്വവുമില്ലാതെ എന്തെങ്കിലുമൊക്കെ ചെയ്യുന്നതാണ്. മലയാളം ടൈപ്പ് ചെയ്യാനുള്ള ഇൻസ്ക്രിപ്റ്റ് രീതിയുടെ പുതുക്കിയ പതിപ്പ് സിഡാക്ക് ഇപ്രകാരം വളരെ തെറ്റായി തയ്യാറാക്കിയ ഒന്നാണ്. സ്വതന്ത്ര മലയാളം കമ്പ്യൂട്ടിങ്ങ് ഇതിലെ തെറ്റുകൾ ചൂണ്ടിക്കാണിച്ചിട്ടും പ്രതികരണമില്ല. മലയാള വെബ്‌വിലാസങ്ങൾക്കായുള്ള IDN സ്റ്റാൻഡേഡും ഇപ്രകാരം അബദ്ധജടിലമായിരുന്നു. ഇതു ചൂണ്ടിക്കാണിച്ചപ്പോൾ വീണ്ടും സ്റ്റാൻഡേഡ് തയ്യാറാക്കാനുള്ള ചർച്ചകൾ തുടങ്ങാൻ സിഡാക്ക് തീരുമാനിച്ചിട്ടുണ്ട്. പക്ഷേ ഇത്തരം സന്ദർഭങ്ങളിലൊന്നും ചുരുക്കം ചില സന്നദ്ധ സാങ്കേതികപ്രവർത്തകരുടെതല്ലാതെ സർക്കാറിന്റെ ഭാഗത്തു നിന്നോ ഭാഷാവിദഗ്ദ്ധർമാരിൽ നിന്നോ പ്രതികരണമുണ്ടാവാറില്ല. ഇതിനുകാരണം ഇവയെപ്പറ്റി പലപ്പോഴും ഇവർ അറിയാതെപ്പോകുന്നതാണ്. ഈ സ്ഥിതി മാറണം. ഐടി വകുപ്പ് ഇക്കാര്യങ്ങളിൽ ഉത്തരവാദിത്തത്തോടെ പ്രവർത്തിക്കേണ്ടിയിരിക്കുന്നു.

സാങ്കേതികവിദ്യാഗവേഷണവും വികസനവും

സ്വതന്ത്ര സോഫ്ററ്‌വെയർ പ്രവർത്തകരുടെ സജീവ പ്രവർത്തനങ്ങളുടെ ഫലമായി മലയാളം ഇന്ത്യൻ ഭാഷകളെ സംബന്ധിച്ചിടത്തോളം സാങ്കേതികവിദ്യയിൽ വളരെ മുന്നിലാണ്. ഏതെങ്കിലും സോഫ്റ്റ്‌വയർ കമ്പനികൾ ഗവേഷണവും സോഫ്റ്റ്‌വെയർ നിർമ്മാണവും നടത്തും , നമ്മൾ ഉപയോഗിക്കും എന്ന നയം മാറണം. സ്വതന്ത്ര സോഫ്റ്റ്‌വയർ വികസനരീതിയിൽ എല്ലാവർക്കും സജീവമായി പങ്കെടുക്കാൻ കഴിയും. ഇക്കാര്യത്തിൽ നമുക്ക് വേണ്ടതു് പുതിയൊരു മാതൃകയുടെ നിർമ്മിതിയാണ്. സർക്കാർ സ്ഥാപനങ്ങളും സ്വതന്ത്ര സോഫ്റ്റ്‌വയർ സമൂഹങ്ങളും തമ്മിലുള്ള പരസ്പര സഹകരണം ഉറപ്പാക്കാനുള്ള നടപടികൾ വേണം. ഇതിന്റെ ആദ്യപടിയായി സർക്കാർ ഏജൻസികൾ അവരുടെ സോഫ്റ്റ്‌വെയർ വികസനരീതി സ്വതന്ത്ര സോഫ്റ്റ്‌വെയർ മാതൃകയിലേക്കു മാറ്റണം. ഇപ്പോളുള്ള സർക്കാർ വക കുത്തക സോഫ്റ്റ്‌വെയറുകൾ ജനങ്ങൾക്കായി സ്വതന്ത്ര ലൈസൻസുകളിൽ തുറന്നു കൊടുക്കണം. ജനങ്ങളുടെ പണമുപയോഗിച്ച് സർക്കാർ ഏജൻസികൾ കുത്തക സോഫ്റ്റ്‌വെയറുണ്ടാക്കി ബൈനറികൾ മാത്രം കൊടുക്കുന്നതു് ഒട്ടും അംഗീകരിക്കാൻ പറ്റില്ല. സ്വതന്ത്ര സോഫ്റ്റ്‌വെയർ ഉപയോഗിച്ച് മലയാളഭാഷക്കു വേണ്ടി ഗവേഷണവും വികസനവും നടത്തുന്നവരെ പ്രോത്സാഹിപ്പിക്കണം. മലയാളികൾക്കു മാത്രമേ കുറ്റമറ്റ രീതിയിൽ മലയാളം സോഫ്റ്റ്‌വെയറുകൾ ഉണ്ടാക്കാൻ കഴിയൂ എന്ന വസ്തുത തിരിച്ചറിയണം. സാങ്കേതികപ്രവർത്തകരും മലയാളഭാഷാ വിദഗ്ദ്ധരും ഒരുമിച്ച് പ്രവർത്തിക്കേണ്ട ഒരു മേഖലയായതിനാൽ ഇവർ തമ്മിലുള്ള സഹകരണം ഉറപ്പാക്കണം. പ്രശ്നങ്ങൾ ചർച്ച ചെയ്യാനും തീരുമാനങ്ങളെടുക്കാനും ഇരുവരെയും പങ്കെടുപ്പിച്ചുള്ള സംവാദങ്ങൾ, ശില്പശാലകൾ എന്നിവ സംഘടിപ്പിക്കണം.

പ്രാദേശികവത്കരിക്കപ്പെട്ട സ്വതന്ത്ര ഓപ്പറേറ്റിങ്ങ് സിസ്റ്റങ്ങൾ സർക്കാറിന്റെ ഇ-ഗവേണൻസ് പദ്ധതികളിൽ ഉപയോഗിക്കണം. പ്രാദേശികവത്കരണത്തിൽ ഇപ്പൊൾ സന്നദ്ധ സാങ്കേതികപ്രവർത്തകർ മാത്രമാണ് പ്രവർത്തിക്കുന്നതു്. ഈ മേഖലയിൽ സ്വതന്ത്ര സോഫ്റ്റ്‌വെയർ കൂട്ടായ്മകളുമായി സർക്കാർ ഏജൻസികൾക്ക് സഹകരിക്കാവുന്നതാണ്.

പരിശീലനം, പ്രചാരണം

മലയാളഭാഷാ കമ്പ്യൂട്ടിങ്ങിൽ ഉപയോക്താക്കൾക്ക് പരിശീലനം നൽകുന്നതിന് സംസ്ഥാനത്തുടനീളം വിദ്യാഭ്യാസസ്ഥാപനങ്ങളോടും സ്വതന്ത്ര സോഫ്റ്റ്‌വെയർ കൂട്ടയ്മകളോടും സഹകരിച്ച് പരിപാടികൾ സംഘടിപ്പിക്കണം. ഉപയോക്തൃപരിശീലനത്തോടൊപ്പം പുതിയ സാങ്കേതിക പ്രവർത്തകരെ ഈ മേഖലയിലേക്ക് ആകർഷിക്കുന്നതിന് സാങ്കേതിക വിദ്യാഭ്യാസ സ്ഥാപനങ്ങളിൽ പരിശീലനപരിപാടികളും

മത്സരങ്ങളും സംഘടിപ്പിക്കണം. ബിരുദ ബിരുദാനന്തര കോഴ്സുകളിൽ മലയാളം കമ്പ്യൂട്ടേഷണൽ ലിംഗ്വിസ്റ്റിക്സ് ഒരു പഠന വിഷയമാക്കണം.

2011 ജനവരി 1,2,3 തിയ്യതികളിൽ തിരുവനന്തപുരത്തു നടന്ന കേരളപഠന കോൺഗ്രസ്സിലെ മലയാളം കമ്പ്യൂട്ടിങ്ങ് എന്ന വിഭാഗത്തിൽ അവതരിപ്പിച്ചത്.

Content licensed under creative commons attribution share alike 3.0