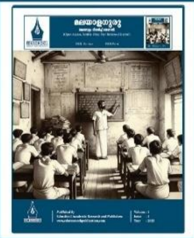
മലയാളഗുരു
(മലയാളം റിസർച്ച് ജേണൽ)
(Open Access, Double-Blind Peer Reviewed Journal)
ISSN Online : 3108-0960 ISSN Print :
EDUSCHOOL
ACADEMIC RESEARCH & PUBLISHERS



ഡിജിറ്റൽ യുഗത്തിലെ മലയാള ഭാഷയുടെ നിലനിൽപ്പും വെല്ലുവിളികളും

Martin Joseph

Assistant Professor, Department of Malayalam, Nirmalagiri College (Autonomous)

Article information

Received: 14th January 2026
Received in revised form: 4th February 2026
Accepted: 3rd March 2026
Available online: 25th April 2026

Volume: 2
Issue: 2
DOI: https://doi.org/10.63090/MG/3108.0960.0026

Abstract

The unprecedented growth of digital technology is significantly impacting the survival of languages around the world. Malayalam, which has been granted classical language status, is going through a complex phase of survival and evolution in the digital age. While Unicode support, social media usage, OTT platforms, and artificial intelligence technology are opening up new possibilities for the Malayalam language, the spread of Manglish, challenges such as language degradation among the younger generation and digital content inequality also persist. This is a study that comprehensively analyzes the status of the Malayalam language in digital media, language preservation efforts, and the potential of NLP technology. This article underscores the importance of policy interventions and social awareness in shaping the future of Malayalam in the digital environment.

സാരാംശം

ഡിജിറ്റൽ സാങ്കേതികവിദ്യയുടെ അഭൂതപൂർവമായ വളർച്ച ലോകമെമ്പാടുമുള്ള ഭാഷകളുടെ നിലനിൽപ്പിനെ സാരമായി സ്വാധീനിക്കുന്നുണ്ട്. ക്ലാസിക്കൽ ഭാഷാ പദവി ലഭിച്ച മലയാളം ഡിജിറ്റൽ യുഗത്തിൽ അതിജീവനത്തിന്റെയും പരിണാമത്തിന്റെയും ഒരു സങ്കീർണമായ ഘട്ടത്തിലൂടെ കടന്നുപോകുകയാണ്. യൂണിക്കോഡ് പിന്തുണ, സോഷ്യൽ മീഡിയ ഉപയോഗം, ഒ.ടി.ടി. പ്ലാറ്റ്ഫോമുകൾ, കൃത്രിമ ബുദ്ധി സാങ്കേതികവിദ്യ എന്നിവ മലയാള ഭാഷയ്ക്ക് പുതിയ സാധ്യതകൾ തുറക്കുമ്പോൾ തന്നെ, മംഗ്ലീഷിന്റെ വ്യാപനം, യുവതലമുറയിലെ ഭാഷാശോഷണം, ഡിജിറ്റൽ ഉള്ളടക്ക അസമത്വം തുടങ്ങിയ വെല്ലുവിളികളും നിലനിൽക്കുന്നു. ഡിജിറ്റൽ മാധ്യമങ്ങളിൽ മലയാള ഭാഷയുടെ സ്ഥിതിവിശേഷം, ഭാഷാ സംരക്ഷണ ശ്രമങ്ങൾ, എൻ.എൽ.പി. സാങ്കേതികവിദ്യയുടെ സാധ്യതകൾ എന്നിവ സമഗ്രമായി വിശകലനം ചെയ്യുന്ന ഒരു പഠനമാണിത്. ഡിജിറ്റൽ പരിസ്ഥിതിയിൽ മലയാളത്തിന്റെ ഭാവി രൂപപ്പെടുത്തുന്നതിൽ നയപരമായ ഇടപെടലുകളുടെയും സാമൂഹിക ബോധവൽക്കരണത്തിന്റെയും പ്രാധാന്യം ഈ ലേഖനം അടിവരയിടുന്നു.

പ്രധാന വാക്കുകൾ:- ഡിജിറ്റൽ മലയാളം, യൂണിക്കോഡ് പിന്തുണ, മംഗ്ലീഷ്, സോഷ്യൽ മീഡിയ ഭാഷാ മാറ്റങ്ങൾ, പ്രകൃതി ഭാഷാ സംസ്കരണം (NLP), ഭാഷാ സംരക്ഷണം

ആമുഖം

ഇരുപത്തിയൊന്നാം നൂറ്റാണ്ടിന്റെ ആദ്യ ദശകങ്ങളിൽ ഡിജിറ്റൽ സാങ്കേതികവിദ്യ മനുഷ്യജീവിതത്തിന്റെ എല്ലാ മേഖലകളിലും കടന്നുകയറിയിരിക്കുന്നു. ഭാഷ ചിന്തയുടെയും സംസ്കാരത്തിന്റെയും ആശയവിനിമയത്തിന്റെയും അടിസ്ഥാന ഉപകരണം ഈ ഡിജിറ്റൽ

വിപ്ലവത്തിൽ ഏറ്റവും ആഴത്തിൽ സ്വാധീനിക്കപ്പെടുന്ന മേഖലകളിലൊന്നാണ്. യുനെസ്കോയുടെ കണക്കനുസരിച്ച്, ലോകത്തെ ഏകദേശം 7,000 ഭാഷകളിൽ പകുതിയിലേറെയും ഈ നൂറ്റാണ്ടിന്റെ അവസാനത്തോടെ ഇല്ലാതാകുമെന്ന് ആശങ്കപ്പെടുന്നുണ്ട്.[1]

മലയാളം, ഏകദേശം നാല് കോടി ആളുകൾ സംസാരിക്കുന്ന, ദ്രാവിഡ ഭാഷാ കുടുംബത്തിലെ ഒരു പ്രധാന ഭാഷയാണ്. 2013-ൽ ഇന്ത്യൻ സർക്കാർ മലയാളത്തിന് ക്ലാസിക്കൽ ഭാഷാ പദവി നൽകിയത് അതിന്റെ സാഹിത്യ-സാംസ്കാരിക പൈതൃകത്തിന്റെ അംഗീകാരമായിരുന്നു. എന്നാൽ ഡിജിറ്റൽ ലോകത്ത് ഈ ഭാഷ നേരിടുന്ന വെല്ലുവിളികൾ നിരവധിയാണ്. ഇംഗ്ലീഷിന്റെ ആധിപത്യം, സാങ്കേതിക പദാവലിയുടെ അഭാവം, ഡിജിറ്റൽ സാക്ഷരതയിലെ അസമത്വം എന്നിവ മലയാളത്തിന്റെ ഡിജിറ്റൽ ഭാവിയെ ചോദ്യം ചെയ്യുന്നു.[2]

ഡിജിറ്റൽ ഭാഷാ പരിസ്ഥിതിയെ (Digital Language Ecology) സംബന്ധിച്ച പഠനങ്ങൾ ഭാഷയുടെ നിലനിൽപ്പിനെ സ്വാധീനിക്കുന്ന ഘടകങ്ങളെ വിശകലനം ചെയ്യുന്നുണ്ട്. ഡിജിറ്റൽ പ്ലാറ്റ്ഫോമുകളിൽ ഒരു ഭാഷയുടെ സാന്നിധ്യം, ആ ഭാഷയിലെ ഡിജിറ്റൽ ഉള്ളടക്കത്തിന്റെ അളവും ഗുണനിലവാരവും, സാങ്കേതിക ഉപകരണങ്ങളുടെ ലഭ്യത, ഉപയോക്താക്കളുടെ ഡിജിറ്റൽ ഭാഷാ ശീലങ്ങൾ എന്നിവ ഈ പാരിസ്ഥിതിക ചട്ടക്കൂടിന്റെ ഭാഗമാണ്. ഈ ലേഖനം ഡിജിറ്റൽ യുഗത്തിലെ മലയാള ഭാഷയുടെ നിലനിൽപ്പിനെയും വെല്ലുവിളികളെയും ഈ ചട്ടക്കൂടിനകത്ത് വിശകലനം ചെയ്യാൻ ശ്രമിക്കുന്നു.[3]

ഡിജിറ്റൽ പ്ലാറ്റ്ഫോമുകളിൽ മലയാളം: ചരിത്രവും വർത്തമാനവും

മലയാളത്തിന്റെ ഡിജിറ്റൽ ചരിത്രം 1990-കളിലെ ഇന്റർനെറ്റ് വ്യാപനത്തോടെ ആരംഭിക്കുന്നു. ആദ്യകാല മലയാളം വെബ്സൈറ്റുകളും ഇ-മെയിൽ ഗ്രൂപ്പുകളും ASCII അടിസ്ഥാനത്തിലുള്ള ഫോണ്ടുകളിലാണ് പ്രവർത്തിച്ചിരുന്നത്, ഇത് വ്യാപക പൊരുത്തക്കേടിനും വായനാ ബുദ്ധിമുട്ടിനും കാരണമായി. 2001-ൽ യൂണിക്കോഡ് 3.1-ൽ മലയാള ലിപി ഉൾപ്പെടുത്തിയത് ഒരു നിർണായക വഴിത്തിരിവായിരുന്നു. ഇതോടെ മലയാളത്തിൽ ഡിജിറ്റൽ ഉള്ളടക്കം സൃഷ്ടിക്കുന്നതിനും പങ്കിടുന്നതിനുമുള്ള സാങ്കേതിക അടിത്തറ ശക്തമായി.[4]

മലയാളം വിക്കിപീഡിയ 2002-ൽ ആരംഭിച്ചു, ഇന്ന് 80,000-ലേറെ ലേഖനങ്ങളുള്ള ഒരു വിജ്ഞാനകോശമായി വളർന്നിരിക്കുന്നു. ബ്ലോഗോസ്ഫിയറിൽ 'ചിന്തിര'യും 'പൂമ്പാറ്റ'യും പോലുള്ള മലയാളം ബ്ലോഗ് അഗ്രഗേറ്ററുകൾ 2005-2010 കാലഘട്ടത്തിൽ മലയാള ഡിജിറ്റൽ എഴുത്തിന്റെ ഒരു നവോത്ഥാനം തന്നെ സൃഷ്ടിച്ചു. ഈ കാലഘട്ടത്തിൽ നൂറുകണക്കിന് മലയാളം ബ്ലോഗുകൾ സാഹിത്യം, രാഷ്ട്രീയം, സംസ്കാരം, സാങ്കേതികവിദ്യ എന്നിവയിൽ സജീവമായി ഉള്ളടക്കം ഉൽപാദിപ്പിച്ചു. എന്നാൽ സോഷ്യൽ മീഡിയയുടെ വരവോടെ ബ്ലോഗ് സംസ്കാരം ക്ഷയിക്കുകയും, കൂടുതൽ ചുരുങ്ങിയ, ചിത്ര-അധിഷ്ഠിത ഉള്ളടക്ക ശൈലികൾ പ്രചാരത്തിലാവുകയും ചെയ്തു.

ഗൂഗിൾ, ഫേസ്ബുക്ക്, വാട്ട്സ്ആപ്പ് തുടങ്ങിയ പ്രധാന ഡിജിറ്റൽ പ്ലാറ്റ്ഫോമുകൾ ഇന്ന് മലയാള ഭാഷാ ഇന്റർഫേസ് നൽകുന്നുണ്ട്. ഗൂഗിൾ ട്രാൻസ്ലേറ്റ്, ഗൂഗിൾ അസിസ്റ്റന്റ്, ഗൂഗിൾ ലെൻസ് എന്നിവയിലെല്ലാം മലയാള പിന്തുണ ലഭ്യമാണ്. എന്നിരുന്നാലും, ഈ പ്ലാറ്റ്ഫോമുകളിലെ മലയാള ഉള്ളടക്കത്തിന്റെ ഗുണനിലവാരം ഇംഗ്ലീഷ് ഉള്ളടക്കത്തെ അപേക്ഷിച്ച് വളരെ താഴ്ന്ന നിലയിലാണ്. സ്വയമേവയുള്ള വിവർത്തന ഗുണനിലവാരം മെച്ചപ്പെടുന്നുണ്ടെങ്കിലും, സങ്കീർണമായ ഭാഷാ പ്രയോഗങ്ങളിലും സാംസ്കാരിക സന്ദർഭങ്ങളിലും ഇപ്പോഴും ഗുരുതരമായ പിഴവുകൾ ഉണ്ടാകുന്നുണ്ട്.[5]

യൂണിക്കോഡും മലയാളം കമ്പ്യൂട്ടിങ്ങും

യൂണിക്കോഡ് കൺസോർഷ്യം മലയാള ലിപിക്ക് U+0D00 മുതൽ U+0D7F വരെയുള്ള കോഡ് പോയിന്റുകൾ നിർണയിച്ചിട്ടുണ്ട്. മലയാളത്തിന്റെ സങ്കീർണമായ ലിപി ഘടന കൂട്ടക്ഷരങ്ങൾ, ചില്ലക്ഷരങ്ങൾ, ചന്ദ്രക്കല, സംവൃതോകാരം എന്നിവ യൂണിക്കോഡ് എൻകോഡിങ്ങിൽ പ്രത്യേക ശ്രദ്ധ ആവശ്യപ്പെട്ടിരുന്നു. 2005-ൽ ചില്ലക്ഷരങ്ങൾക്ക് സ്വതന്ത്ര കോഡ് പോയിന്റുകൾ നൽകുന്നതു സംബന്ധിച്ച തർക്കം മലയാളം കമ്പ്യൂട്ടിങ്ങ് ചരിത്രത്തിലെ ഒരു പ്രധാന അധ്യായമായിരുന്നു.[6]

സ്വതന്ത്ര സോഫ്റ്റ്‌വെയർ പ്രസ്ഥാനം മലയാളം കമ്പ്യൂട്ടിങ്ങിന്റെ വളർച്ചയിൽ നിർണായക പങ്ക് വഹിച്ചിട്ടുണ്ട്. സ്വതന്ത്ര മലയാളം കമ്പ്യൂട്ടിങ്ങ് (SMC) എന്ന സന്നദ്ധ സംഘടന മലയാള ഫോണ്ടുകൾ, ഇൻപുട്ട് മെത്തഡുകൾ, മോർഫോളജിക്കൽ അനലൈസർ, സ്പെൽ ചെക്കർ തുടങ്ങിയ ഉപകരണങ്ങൾ വികസിപ്പിച്ചെടുത്തിട്ടുണ്ട്. രചന, മീര, AnjaliOldLipi തുടങ്ങിയ

യൂണിക്കോഡ് ഫോണ്ടുകൾ മലയാള ഡിജിറ്റൽ എഴുത്തിന്റെ വ്യാപനത്തിന് വഴിയൊരുക്കി. ലിബ്രെഓഫീസിലും ലിനക്സ് ഡിസ്ട്രിബ്യൂഷനുകളിലും മലയാള പിന്തുണ ശക്തമാക്കുന്നതിൽ ഈ സമൂഹത്തിന്റെ സംഭാവന അമൂല്യമാണ്.[7]

മൊബൈൽ ഫോണുകളിലെ മലയാള ഇൻപുട്ട് രീതികളുടെ പരിണാമം ഭാഷയുടെ ഡിജിറ്റൽ ഉപയോഗത്തെ ഗണ്യമായി സ്വാധീനിച്ചിട്ടുണ്ട്. ആദ്യകാല ഫോണുകളിൽ മലയാള ഇൻപുട്ട് അസാധ്യമായിരുന്നു, ഇത് മംഗ്ലീഷിന്റെ (ലാറ്റിൻ ലിപിയിൽ മലയാളം എഴുതുന്ന ശൈലി) വ്യാപനത്തിന് കാരണമായി. ഇന്ന് ഗൂഗിൾ ഇൻഡിക് കീബോർഡ്, സ്വിഫ്റ്റ്കീ, വരമൊഴി തുടങ്ങിയ ഇൻപുട്ട് ഉപകരണങ്ങൾ ട്രാൻസ്ലിറ്ററേഷൻ, ഹാൻഡ്റൈറ്റിങ്, വോയ്സ് ഇൻപുട്ട് എന്നിവ വഴി മലയാളം ടൈപ്പിംഗ് എളുപ്പമാക്കിയിട്ടുണ്ട്. എന്നിട്ടും, പലരും ഇപ്പോഴും മംഗ്ലീഷ് തിരഞ്ഞെടുക്കുന്നു ഇത് ശീലത്തിന്റെയും വേഗതയുടെയും പ്രശ്നമാണ്.

സോഷ്യൽ മീഡിയയും മലയാള ഭാഷാ മാറ്റങ്ങളും

സോഷ്യൽ മീഡിയ പ്ലാറ്റ്ഫോമുകൾ മലയാള ഭാഷയുടെ ഉപയോഗ ശൈലിയിൽ സമൂലമായ മാറ്റങ്ങൾ വരുത്തിയിട്ടുണ്ട്. ഫേസ്ബുക്ക്, ട്വിറ്റർ (ഇപ്പോൾ എക്സ്), ഇൻസ്റ്റഗ്രാം എന്നിവയിലെ മലയാള ഭാഷാ ഉപയോഗം പഠനവിധേയമാക്കിയ ഗവേഷണങ്ങൾ ചില ശ്രദ്ധേയമായ പ്രവണതകൾ കണ്ടെത്തിയിട്ടുണ്ട്. കോഡ്-മിക്സിങ് (code-mixing), അതായത് ഒരേ വാക്യത്തിൽ മലയാളവും ഇംഗ്ലീഷും കലർത്തി ഉപയോഗിക്കുന്ന ശൈലി, സോഷ്യൽ മീഡിയ ഭാഷയുടെ ഒരു മുഖ്യ സവിശേഷതയായി മാറിയിരിക്കുന്നു.[8]

ട്വിറ്ററിലെ മലയാള ട്വീറ്റുകളുടെ ഒരു പഠനം കാണിക്കുന്നത്, ഏകദേശം 40 ശതമാനം ട്വീറ്റുകൾ മലയാളം-ഇംഗ്ലീഷ് കോഡ്-മിക്സിങ് ശൈലിയിലാണ് എന്നാണ്. ഇത് ഭാഷാശാസ്ത്ര ദൃഷ്ടിയിൽ ഒരു പ്രതിഭാസമാണ് ഡിഗ്ലോസിയയുടെ (diglossia) ഒരു പുതിയ രൂപം. ഔപചാരിക എഴുത്തിൽ ശുദ്ധ മലയാളവും, അനൗപചാരിക ഡിജിറ്റൽ ആശയവിനിമയത്തിൽ കോഡ്-മിക്സ്ഡ് ഭാഷയും ഉപയോഗിക്കുന്ന ഈ പ്രവണത ഭാഷയുടെ പരിണാമത്തിന്റെ ഭാഗമാണോ, അതോ ക്ഷയത്തിന്റെ ലക്ഷണമാണോ എന്ന ചോദ്യം ഗൗരവമേറിയതാണ്.

സോഷ്യൽ മീഡിയയിൽ മലയാളം ലിപിയുടെ ഉപയോഗം വർധിച്ചുവരുന്നു എന്നത് ഒരു ശുഭകരമായ പ്രവണതയാണ്. 2015-നു മുൻപ് ഫേസ്ബുക്കിലെ മലയാള ഉള്ളടക്കത്തിന്റെ ഭൂരിഭാഗവും മംഗ്ലീഷിലായിരുന്നു. എന്നാൽ സ്മാർട്ട്ഫോണുകളിലെ മലയാള കീബോർഡ് സാങ്കേതികവിദ്യയുടെ മെച്ചപ്പെടലിനും ഭാഷാ ബോധവൽക്കരണ പ്രവർത്തനങ്ങൾക്കും ശേഷം, മലയാള ലിപിയിലെ പോസ്റ്റുകളുടെ എണ്ണം ഗണ്യമായി വർധിച്ചിട്ടുണ്ട്. 'ഫേസ്ബുക്കിൽ മലയാളം' പോലുള്ള കൂട്ടായ്മകൾ ഈ മാറ്റത്തിന് ഗണ്യമായ സംഭാവന നൽകിയിട്ടുണ്ട്.[9]

മീം സംസ്കാരം മലയാള ഭാഷയ്ക്ക് ഡിജിറ്റൽ ലോകത്ത് പ്രത്യേകമായ ഒരിടം ഉറപ്പാക്കിയിട്ടുണ്ട്. മലയാള സിനിമാ ഡയലോഗുകൾ അടിസ്ഥാനമാക്കിയുള്ള മീമുകൾ ഇന്ത്യയിലെ ഇന്റർനെറ്റ് സംസ്കാരത്തിൽ ഒരു സവിശേഷ സ്ഥാനം നേടിയിട്ടുണ്ട്. 'ട്രോൾ മലയാളം', 'ഇന്റർനാഷണൽ ചളു യൂണിയൻ' തുടങ്ങിയ ഗ്രൂപ്പുകൾ കോടിക്കണക്കിന് ഫോളോവേഴ്സുമായി മലയാള ഭാഷയുടെ ഡിജിറ്റൽ സാന്നിധ്യം ശക്തിപ്പെടുത്തുന്നുണ്ട്. എന്നാൽ ഈ ഉള്ളടക്കത്തിൽ ഭാഷയുടെ ലളിതവൽക്കരണവും വ്യാകരണ നിയമങ്ങളുടെ അവഗണനയും സംഭവിക്കുന്നുണ്ട് എന്നത് ശ്രദ്ധിക്കേണ്ട വസ്തുതയാണ്.

ഒ.ടി.ടി. പ്ലാറ്റ്ഫോമുകളും മലയാള ഡിജിറ്റൽ ഉള്ളടക്കവും

ഒ.ടി.ടി. (Over-The-Top) സ്ട്രീമിങ് പ്ലാറ്റ്ഫോമുകളുടെ വരവ് മലയാള ഭാഷാ ഉള്ളടക്കത്തിന്റെ ആഗോള വ്യാപനത്തിന് പുതിയ വാതിലുകൾ തുറന്നിട്ടുണ്ട്. നെറ്റ്ഫ്ലിക്സ്, ആമസോൺ പ്രൈം, ഹോട്ട്സ്റ്റാർ, സോണി ലിവ്, മനോരമ മാക്സ് തുടങ്ങിയ പ്ലാറ്റ്ഫോമുകൾ മലയാള സിനിമകളും ഒറിജിനൽ സീരീസുകളും ലോകമെമ്പാടുമുള്ള പ്രേക്ഷകർക്ക് ലഭ്യമാക്കുന്നു. ഇത് മലയാള ഭാഷയുടെ സാംസ്കാരിക ആഗോള സാന്നിധ്യം ശക്തിപ്പെടുത്തുന്നതിൽ നിർണായക പങ്ക് വഹിക്കുന്നുണ്ട്.[10]

യൂട്യൂബ് മലയാള ഭാഷാ ഉള്ളടക്ക സൃഷ്ടിക്ക് ഏറ്റവും പ്രധാനപ്പെട്ട പ്ലാറ്റ്ഫോമാണ്. വിദ്യാഭ്യാസം, പാചകം, യാത്ര, സാങ്കേതികവിദ്യ, വാർത്തകൾ, സിനിമാ നിരൂപണം എന്നിങ്ങനെ വിവിധ വിഭാഗങ്ങളിൽ ആയിരക്കണക്കിന് മലയാളം യൂട്യൂബ് ചാനലുകൾ സജീവമാണ്. 'ടെക് ട്രാവലർ', 'വെജിറ്റേറിയൻ ഫുഡ്സ്", 'ഉണ്ണിമായ', 'മലയാളം ടെക്" തുടങ്ങിയ ചാനലുകൾ ലക്ഷക്കണക്കിന് സബ്സ്ക്രൈബേഴ്സുമായി ഡിജിറ്റൽ മലയാളം ഉള്ളടക്ക വ്യവസായത്തെ ശക്തിപ്പെടുത്തുന്നു.

പോഡ്കാസ്റ്റുകളും ഓഡിയോ ഉള്ളടക്ക പ്ലാറ്റ്ഫോമുകളും മലയാള ഭാഷയ്ക്ക് പുതിയ മാധ്യമങ്ങൾ നൽകുന്നുണ്ട്. സ്പോട്ടിഫൈ, ആപ്പിൾ പോഡ്കാസ്റ്റ്, ഗൂഗിൾ പോഡ്കാസ്റ്റ് എന്നിവയിൽ നിരവധി മലയാളം പോഡ്കാസ്റ്റുകൾ ലഭ്യമാണ്. 'കുസൃതി', 'ദ ഫൗർത്ത്', 'സ്റ്റോറി ടെല്ലർ' തുടങ്ങിയ പോഡ്കാസ്റ്റുകൾ മലയാളത്തിൽ ഗുണനിലവാരമുള്ള ഓഡിയോ ഉള്ളടക്കം സൃഷ്ടിക്കുന്നതിൽ മുന്നേറ്റം നടത്തിക്കൊണ്ടിരിക്കുന്നു. എന്നാൽ ഇംഗ്ലീഷ് പോഡ്കാസ്റ്റുകളുമായി താരതമ്യം ചെയ്യുമ്പോൾ, മലയാള ഓഡിയോ ഉള്ളടക്കത്തിന്റെ വൈവിധ്യവും തൊഴിൽപരമായ നിലവാരവും ഇനിയും ഏറെ മെച്ചപ്പെടേണ്ടതുണ്ട്.

യുവതലമുറയിലെ ഭാഷാശോഷണവും ലിപി-മംഗ്ലീഷ് ചർച്ചയും

കേരളത്തിലെ യുവതലമുറയുടെ ഭാഷാ ശീലങ്ങളിൽ ഗണ്യമായ മാറ്റങ്ങൾ സംഭവിച്ചുകൊണ്ടിരിക്കുന്നു. ഇംഗ്ലീഷ് മീഡിയം വിദ്യാഭ്യാസത്തിന്റെ വ്യാപനം, ആഗോളവൽക്കരണം, ഡിജിറ്റൽ മാധ്യമങ്ങളുടെ ഇംഗ്ലീഷ് ആധിപത്യം എന്നിവ ചേർന്ന് മലയാളത്തിന്റെ ഉപയോഗം കുറയ്ക്കുന്നു. 'ഭാഷാ നഷ്ടം' (language attrition) എന്ന ഭാഷാശാസ്ത്ര പ്രതിഭാസം കേരളത്തിലെ നഗര യുവതലമുറയിൽ പ്രകടമാകുന്നുണ്ട്. പല യുവജനങ്ങൾക്കും മലയാളം വായിക്കാനോ എഴുതാനോ ബുദ്ധിമുട്ടുണ്ട്, സംസാരിക്കാൻ കഴിയുമെങ്കിലും.[11]

മംഗ്ലീഷ് മലയാളം ശബ്ദങ്ങൾ ലാറ്റിൻ (ഇംഗ്ലീഷ്) അക്ഷരമാലയിൽ എഴുതുന്ന ശൈലി ഡിജിറ്റൽ കാലത്തെ ഏറ്റവും വിവാദപരമായ ഭാഷാ പ്രതിഭാസങ്ങളിലൊന്നാണ്. മംഗ്ലീഷിനെ അനുകൂലിക്കുന്നവർ വാദിക്കുന്നത്, ഇത് ഭാഷയുടെ പ്രായോഗിക ഉപയോഗം എളുപ്പമാക്കുന്നു, ഡിജിറ്റൽ ആശയവിനിമയത്തിന്റെ വേഗത വർധിപ്പിക്കുന്നു, മലയാള ലിപി അറിയാത്ത പ്രവാസി മലയാളികളെ ഭാഷയുമായി ബന്ധിപ്പിക്കുന്നു എന്നൊക്കെയാണ്. എതിർപക്ഷം ചൂണ്ടിക്കാട്ടുന്നത്, മംഗ്ലീഷ് മലയാള ലിപിയുടെ ക്ഷയത്തിന് വഴിയൊരുക്കുന്നു, ഭാഷയുടെ ദൃശ്യ സൗന്ദര്യവും സാംസ്കാരിക സ്വത്വവും നശിപ്പിക്കുന്നു എന്നാണ്.

ഈ ചർച്ചയിൽ ഒരു മധ്യമ നിലപാട് രൂപപ്പെടുത്തേണ്ടതുണ്ട്. മംഗ്ലീഷ് ഒരു താൽക്കാലിക ആവശ്യമാണ്, എന്നാൽ മലയാള ലിപിയിലെ ഡിജിറ്റൽ സാക്ഷരത വർധിപ്പിക്കുന്നതിന് ബോധപൂർവമായ ശ്രമങ്ങൾ ആവശ്യമാണ്. വിദ്യാഭ്യാസ സ്ഥാപനങ്ങളിൽ മലയാള ലിപിയിലെ ഡിജിറ്റൽ ടൈപ്പിംഗ് പരിശീലനം, മലയാള ലിപി ഇൻപുട്ട് ഉപകരണങ്ങളുടെ ഉപയോഗ സൗകര്യം മെച്ചപ്പെടുത്തൽ, മലയാള ലിപിയിലെ ഉള്ളടക്ക സൃഷ്ടിയെ പ്രോത്സാഹിപ്പിക്കൽ തുടങ്ങിയ നടപടികൾ അനിവാര്യമാണ്. കേരള സർക്കാരിന്റെ ഐ.ടി. നയത്തിൽ മലയാള ഭാഷാ ഡിജിറ്റൽ സാക്ഷരത ഒരു പ്രധാന ഘടകമായി ഉൾപ്പെടുത്തേണ്ടത് അത്യാവശ്യമാണ്.

ഡിജിറ്റൽ സംരക്ഷണ ശ്രമങ്ങളും ഭാഷാ നയവും

മലയാള ഭാഷയുടെ ഡിജിറ്റൽ സംരക്ഷണത്തിനായി സർക്കാർ-സ്വകാര്യ തലങ്ങളിൽ വിവിധ ശ്രമങ്ങൾ നടക്കുന്നുണ്ട്. കേരള സർക്കാരിന്റെ ഭാഷാ ഇൻസ്റ്റിറ്റ്യൂട്ട്, ഇന്റർനാഷണൽ സെന്റർ ഫോർ കേരള സ്റ്റഡീസ് (ICKS), ഇന്ത്യൻ ഇൻസ്റ്റിറ്റ്യൂട്ട് ഓഫ് ഇൻഫർമേഷൻ ടെക്നോളജി ആൻഡ് മാനേജ്മെന്റ് - കേരള (IIITM-K) തുടങ്ങിയ സ്ഥാപനങ്ങൾ മലയാളം കമ്പ്യൂട്ടിങ്ങ്, ഭാഷാ സാങ്കേതികവിദ്യ, ഡിജിറ്റൽ പൈതൃക സംരക്ഷണം എന്നിവയിൽ പ്രവർത്തിക്കുന്നു.[12]

സയൻസ് ആൻഡ് ടെക്നോളജി പദാവലി നിർമാണം മലയാള ഭാഷയുടെ ഡിജിറ്റൽ കാലത്തെ നിലനിൽപ്പിന് അനിവാര്യമാണ്. കേരള ഭാഷാ ഇൻസ്റ്റിറ്റ്യൂട്ടിന്റെ ശബ്ദതാരാവലി ഡിജിറ്റൽ പദ്ധതി, ഓൺലൈൻ നിഘണ്ടുക്കൾ, സാങ്കേതിക പദകോശങ്ങൾ എന്നിവ ഈ ദിശയിലുള്ള ശ്രമങ്ങളാണ്. എന്നാൽ സാങ്കേതിക മേഖലയിലെ ദ്രുതഗതിയിലുള്ള മാറ്റങ്ങൾക്കൊപ്പം പദാവലി നവീകരണം നടക്കുന്നില്ല എന്ന വിമർശനം ന്യായമാണ്. ആർട്ടിഫിഷ്യൽ ഇന്റലിജൻസ്, ബ്ലോക്ക്ചെയിൻ, ക്ലൗഡ് കമ്പ്യൂട്ടിങ്ങ്, ക്വാണ്ടം കമ്പ്യൂട്ടിങ്ങ് തുടങ്ങിയ മേഖലകളിലെ മലയാള സാങ്കേതിക പദാവലി ഇപ്പോഴും അപര്യാപ്തമാണ്.

ഡിജിറ്റൽ പൈതൃക സംരക്ഷണ പദ്ധതികൾ മലയാള ഭാഷയുടെ ചരിത്ര രേഖകൾ ഡിജിറ്റൽ രൂപത്തിൽ സംരക്ഷിക്കുന്നതിൽ പ്രധാന പങ്ക് വഹിക്കുന്നു. പ്രോജക്ട് 'സായാഹ്നം' പഴയ മലയാള ഗ്രന്ഥങ്ങൾ ഡിജിറ്റൈസ് ചെയ്യുന്ന ഒരു പ്രധാന സന്നദ്ധ പദ്ധതിയാണ്. നാഷണൽ ഡിജിറ്റൽ ലൈബ്രറി ഓഫ് ഇന്ത്യ, TDIL (Technology Development for Indian Languages) പ്രോഗ്രാം എന്നിവ കേന്ദ്ര സർക്കാർ തലത്തിലെ ശ്രമങ്ങളാണ്. കേരള സർക്കാരിന്റെ 'ഗ്രന്ഥശാല' ഡിജിറ്റൽ ലൈബ്രറി പദ്ധതി മലയാള പുസ്തകങ്ങളുടെ ഡിജിറ്റൽ ലഭ്യത വർധിപ്പിക്കുന്നതിൽ ഗണ്യമായ സംഭാവന നൽകിയിട്ടുണ്ട്.

കൃത്രിമ ബുദ്ധി, പ്രകൃതി ഭാഷാ സംസ്കരണം, മലയാളത്തിന്റെ ഭാവി

കൃത്രിമ ബുദ്ധി (Artificial Intelligence) സാങ്കേതികവിദ്യയുടെ മുന്നേറ്റം മലയാള ഭാഷയ്ക്ക് പുതിയ സാധ്യതകളും വെല്ലുവിളികളും ഒരേ സമയം സൃഷ്ടിക്കുന്നുണ്ട്. പ്രകൃതി ഭാഷാ സംസ്കരണം (Natural Language Processing - NLP) മേഖലയിൽ മലയാളത്തിനായുള്ള ഗവേഷണം പുരോഗമിക്കുന്നുണ്ടെങ്കിലും, ഇംഗ്ലീഷ് പോലുള്ള ഭാഷകളുമായി താരതമ്യം ചെയ്യുമ്പോൾ ഇനിയും വളരെ ദൂരം സഞ്ചരിക്കാനുണ്ട്. മലയാളം ഇപ്പോഴും ഒരു 'low-resource language' ആയി NLP ഗവേഷണത്തിൽ കണക്കാക്കപ്പെടുന്നു.[13]

മലയാള NLP-യിലെ പ്രധാന വെല്ലുവിളികൾ ഇവയാണ്: ഒന്ന്, ഭാഷാ ഡാറ്റയുടെ അഭാവം വലിയ, ഗുണനിലവാരമുള്ള, ലേബൽ ചെയ്ത കോർപ്പസുകൾ മലയാളത്തിൽ കുറവാണ്. രണ്ട്, ഭാഷയുടെ രൂപശാസ്ത്ര സങ്കീർണത മലയാളം ഒരു അഗ്ലൂട്ടിനേറ്റീവ് ഭാഷയാണ്, ഒരു വാക്കിൽ നിരവധി പ്രത്യയങ്ങൾ ചേർത്ത് അർഥ വ്യത്യാസങ്ങൾ സൃഷ്ടിക്കാം, ഇത് ടോക്കണൈസേഷനും മോർഫോളജിക്കൽ അനാലിസിസും ബുദ്ധിമുട്ടാക്കുന്നു. മൂന്ന്, സ്ക്രിപ്റ്റ് സങ്കീർണത കൂട്ടക്ഷരങ്ങളും ലിഗേച്ചറുകളും ഒ.സി.ആർ. (Optical Character Recognition) സാങ്കേതികവിദ്യയ്ക്ക് വെല്ലുവിളി ഉയർത്തുന്നു.

എന്നിരുന്നാലും, ചില ശ്രദ്ധേയമായ മുന്നേറ്റങ്ങൾ നടക്കുന്നുണ്ട്. ട്രാൻസ്ഫോർമർ അടിസ്ഥാനമായ മൾട്ടിലിങ്ഗ്വൽ മോഡലുകൾ XLM-RoBERTa, mBERT, IndicBERT തുടങ്ങിയവ മലയാളം ഉൾപ്പെടെയുള്ള ഇന്ത്യൻ ഭാഷകൾക്ക് മെച്ചപ്പെട്ട NLP ഫലങ്ങൾ നൽകുന്നുണ്ട്. AI4Bharat-ന്റെ IndicNLP പദ്ധതി, മൈക്രോസോഫ്റ്റിന്റെ Project Turing, ഗൂഗിളിന്റെ MuRIL (Multilingual Representations for Indian Languages) മോഡൽ എന്നിവ ഇന്ത്യൻ ഭാഷകൾക്കായുള്ള NLP ഗവേഷണത്തിൽ കുതിച്ചുചാട്ടം നടത്തിയിട്ടുണ്ട്.[14]

ചാറ്റ്ജി.പി.ടി., ഗൂഗിൾ ജെമിനി തുടങ്ങിയ വൻ ഭാഷാ മോഡലുകൾ (Large Language Models - LLMs) മലയാളത്തിൽ ഒരു പരിധിവരെ പ്രവർത്തിക്കുന്നുണ്ടെങ്കിലും, ഗുണനിലവാരം ഇനിയും മെച്ചപ്പെടേണ്ടതുണ്ട്. മലയാളത്തിലെ സാംസ്കാരിക സൂക്ഷ്മതകൾ, ഭാഷാ പ്രത്യേകതകൾ, സന്ദർഭ ബോധം എന്നിവ ഈ മോഡലുകളിൽ പലപ്പോഴും അപര്യാപ്തമാണ്. മലയാളത്തിനായുള്ള സമർപ്പിത ഭാഷാ മോഡലുകളുടെ വികസനം കൂടുതൽ മലയാള ഡാറ്റ ഉൾപ്പെടുത്തിയ പരിശീലനത്തിലൂടെ ഭാവിയിലെ ഒരു പ്രധാന ആവശ്യമാണ്. കൃത്രിമ ബുദ്ധി സാങ്കേതികവിദ്യ മലയാള ഭാഷയുടെ നിലനിൽപ്പിനെ സഹായിക്കുമോ, അതോ ഇംഗ്ലീഷിന്റെ ആധിപത്യം ശക്തിപ്പെടുത്തുമോ എന്ന ചോദ്യം ഡിജിറ്റൽ ഭാഷാ നയത്തിന്റെ ഭാവി ദിശ നിർണയിക്കുന്ന ഒന്നാണ്.

ഉപസംഹാരം

ഡിജിറ്റൽ യുഗത്തിൽ മലയാള ഭാഷ ഒരേ സമയം സാധ്യതകളുടെയും വെല്ലുവിളികളുടെയും ഒരു സങ്കീർണമായ ഭൂപ്രദേശത്ത് നിലകൊള്ളുന്നു. യൂണിക്കോഡ് പിന്തുണ, ഒ.ടി.ടി. പ്ലാറ്റ്ഫോമുകൾ, സോഷ്യൽ മീഡിയ, AI/NLP സാങ്കേതികവിദ്യ എന്നിവ മലയാളത്തിന്റെ ഡിജിറ്റൽ സാന്നിധ്യം ശക്തിപ്പെടുത്തുന്നു. എന്നാൽ മംഗ്ലീഷിന്റെ വ്യാപനം, യുവതലമുറയിലെ ലിപി സാക്ഷരതയുടെ കുറവ്, NLP ഗവേഷണത്തിലെ ഡാറ്റ അഭാവം, സാങ്കേതിക പദാവലിയുടെ അപര്യാപ്തത എന്നിവ ഗൗരവമേറിയ വെല്ലുവിളികളാണ്.

ഈ വെല്ലുവിളികളെ നേരിടാൻ ഒരു ബഹുമുഖ സമീപനം ആവശ്യമാണ്. ഒന്ന്, വിദ്യാഭ്യാസ മേഖലയിൽ മലയാള ലിപിയിലെ ഡിജിറ്റൽ സാക്ഷരത പരിശീലനം നിർബന്ധമാക്കണം. രണ്ട്, മലയാള NLP ഗവേഷണത്തിന് കൂടുതൽ ധനസഹായവും സ്ഥാപനപരമായ പിന്തുണയും നൽകണം. മൂന്ന്, ഡിജിറ്റൽ പ്ലാറ്റ്ഫോമുകളിൽ മലയാള ഉള്ളടക്ക സൃഷ്ടിയെ പ്രോത്സാഹിപ്പിക്കുന്ന നയപരമായ ഇടപെടലുകൾ ആവശ്യമാണ്. നാല്, സാങ്കേതിക പദാവലി നിർമാണം ത്വരിതപ്പെടുത്തണം. അഞ്ച്, ഡിജിറ്റൽ പൈതൃക സംരക്ഷണ പദ്ധതികൾ ശക്തിപ്പെടുത്തണം.

ഒരു ഭാഷ ഡിജിറ്റൽ ലോകത്ത് നിലനിൽക്കണമെങ്കിൽ അത് ആ ലോകത്ത് ജീവിക്കുന്നവർക്ക് ഉപയോഗപ്രദവും ആകർഷകവും ആയിരിക്കണം. മലയാളത്തിന്റെ സമ്പന്നമായ സാഹിത്യ-സാംസ്കാരിക പൈതൃകം ഡിജിറ്റൽ യുഗത്തിൽ പുനർജനിക്കണമെങ്കിൽ, ഭാഷാ സാങ്കേതികവിദ്യയിലെ നിക്ഷേപവും സാമൂഹിക ബോധവൽക്കരണവും ഒരുമിച്ച് പോകണം. ഡിജിറ്റൽ മലയാളം ഒരു 'ജീവിക്കുന്ന ഭാഷ' ആയി നിലനിർത്തുന്നതിന് ഓരോ മലയാളിക്കും ഉത്തരവാദിത്തമുണ്ട് ഡിജിറ്റൽ ലോകത്ത്

മലയാളത്തിൽ വായിക്കുക, എഴുതുക, സൃഷ്ടിക്കുക എന്നത് ഭാഷാ സംരക്ഷണത്തിന്റെ ഏറ്റവും ഫലപ്രദമായ മാർഗമാണ്.

അവലംബങ്ങൾ

1. UNESCO. *World Atlas of Languages*. 2024. https://en.wal.unesco.org/.
2. Government of India. *Notification: Classical Language Status for Malayalam*. Ministry of Culture. August 8, 2013.
3. Mark Warschauer. *Technology and Social Inclusion: Rethinking the Digital Divide*. Cambridge, MA: MIT Press, 2004.
4. The Unicode Consortium. *The Unicode Standard, Version 15.0: South Asian Scripts — Malayalam*. 2022. https://www.unicode.org/versions/Unicode15.0.0/.
5. Shobha Warrier. “How Google Is Making Its Products Accessible in Indian Languages.” Rediff.com, March 15, 2022.
6. Santhosh Thottingal. “Malayalam Encoding and Rendering in Digital Platforms.” Swathanthra Malayalam Computing, 2011. https://smc.org.in/.
7. Swathanthra Malayalam Computing. “About SMC: Projects and Contributions.” https://smc.org.in/.
8. Kavya Manohar, and Soman K. P.. “An Effective Approach to Morphological Analysis of Malayalam Words for Information Retrieval Applications.” *International Journal of Dravidian Linguistics* 47, no. 2 (2018): 67–82.
9. J. Devika. “Surviving in a Knowledge Economy: Languages and Social Transformations in Kerala.” *Indian Journal of Gender Studies* 17, no. 2 (2010): 197–220.
10. KPMG and Google. *Indian Languages: Defining India's Internet*. April 2017. https://assets.kpmg.com/content/dam/kpmg/in/pdf/2017/04/Indian-languages-Defining-Indias-Internet.pdf.
11. A. Jayashree. “Language Maintenance and Shift in Kerala: A Sociolinguistic Study.” *Language in India* 14, no. 8 (2014): 34–45.
12. IIITM-Kerala. “Centre for Malayalam Language Computing.” https://www.iiitmk.ac.in/.
13. Anoop Kunchukuttan, Divyanshu Kakwani, Satish Golla, et al. “AI4Bharat-IndicNLP Corpus: Monolingual Corpora and Word Embeddings for Indic Languages.” In *Findings of ACL: EMNLP 2020*, 4948–4961, 2020.
14. Simran Khanuja, Diksha Bansal, Sarvesh Mehtani, et al. “MuRIL: Multilingual Representations for Indian Languages.” arXiv preprint, 2021. https://arxiv.org/abs/2103.10730.