

ଓଡ଼ିଆ:ଟୁଲ/ଏନକୋଡିଙ୍ଗ କନଭର୍ଟର

<https://or.wikipedia.org/s/t33>

- ପୃଥକ ପୃଷ୍ଠା
- ସଂପଦ ପୃଷ୍ଠା ଫଳକ
- ନଗଦ ବଦଳ
- ଯାଜ୍ଞିତାହି ପୃଷ୍ଠା
- ଦାନ
- ପୋଷାପୋଷ
- ତାଟପତା
- ଆଲୋଚନା ସଭା
- ସହଯୋଗ କରନ୍ତୁ
- ଓଡ଼ିଆ ଟୁଲ୍‌କିଟ୍
- ଦକ୍ଷା ଛବି
- ଦକ୍ଷା ଲେଖା
- ଉପକରଣ
- ଏଠାରେ ଥିବା ଲିଙ୍କ
- ପଞ୍ଜିକୃତ ବଦଳ
- ଫାଇଲ ଅପଲୋଡ୍ କରିବେ
- ବିଶେଷ ପୃଷ୍ଠା
- ସବୁବିଧିଆ ଲିଙ୍କ
- ପୃଷ୍ଠା ସୂଚନା
- କ୍ଷୋର URL
- ପୃଷ୍ଠା/ଉପାଳି
- ବହି ତିଆରି କରିବେ
- PDF ଭାବରେ ଡାଉନଲୋଡ୍ କରିବେ
- ଛପାହୋଇପାରିବା ପୃଷ୍ଠା
- Languages
- Add links

How we built Odia-script encoding converters and expand Wikimedia projects -Jnanaranjan Sahu

Shortcut:
WP:Converter



ଅ*ଠା ଅଠRèkò; Kcòà Lyò: ieKüè... ଚକ୍ରିବ... ଶି; ଶିବ ଶର୍ଦ୍ଦ: ସରକାରୀ ଅଫିସରୁ ଆରମ୍ଭ

Unicode Characters



- Unicode is a computing industry standard
- Covers almost all languages
- Used in all modern industries
- Lots of freely-licensed fonts available

Non Unicode Characters

- ASCII and ISCII
- Outdated
- Not searchable
- Maximum fonts are proprietary

How to deal with encoding mismatch?

- Type everything again in Unicode
- Create a converter that converts the non-Unicode encoding to Unicode
- Use OCR to convert encoding to Unicode

Building blocks of converters for Odia

Rebati Converter by Srujanika

Rebati Converter 2

AkrutiOriSarala99 TO Unicode Odia Font Converter

AkrutiOriSarala99 font text-box

27 G[e aō Mēēāūēē Cēē ^ōēē.ūjR^Kō
28 zēZūd [e _ūkk UōūK ējōK icd _ūAñ MēēāūēēKē _ūYō bōZūē cūWYō aīōūf, ēūC zū' _ūē CVūA ēūYō _ēYō
Kjōūf- G[e ēagY ūbu ūjūA[ōāō
29 Z[ū ō Mēēāūēē Cēē icū^ [ōfūō 30 GYē ēūC UōūK ējōK icd _ūAñ zū' cēōKē ūYō bōZūē cūWYō\ūf
_ūkkō ūZūā aō ābē -ūgēū \gō^ _ūA^ [ōfū Mēēāūēē gaeō 31 GjōbKō ēūV [e ūjfuō āZō[e Wēā^ē icd
axYō[ōGō Z[ū ō Mēēāūēē ūīA cKū Cēēō 32 aōg_b Cēē [ūA _ūPōKūf, 33 'Gjūē ū_būē axYō-ūAQōō zū'e
icē ū ūjōZ ūjūā -ōY^ zūKē Wēā^ \ōēō ēūc Pūjēñ zū'e _ēōZāūYō ūi ū-ūjZē LāūAūē aōgYūi eLōQō, ēagY
_ābē-ūgē zūKē icū ūūāō'

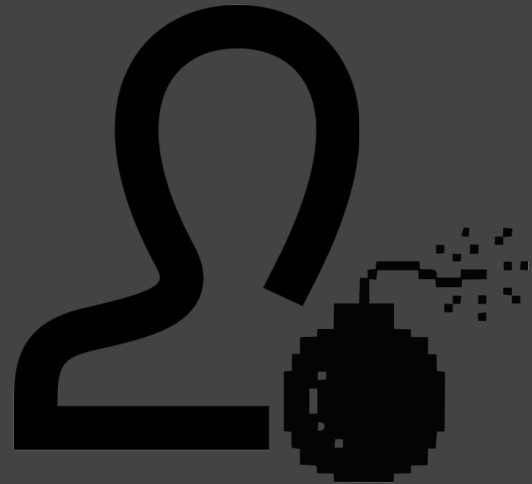
Convert to Unicode >>

Unicode text-box

ଯାଶୁକ ସନ ସାଧନ
୧୫ ବିଶେଷ ଦୁନଶ୍ଚ କହିବା ଆରମ୍ଭ କଲେ। ୧୬ 'ସମସ୍ତେ ପାପ ପଥକୁ ଫେରା କାନ୍ଦଣ ସୁନିରାଜ୍ୟ ନିକଟତର। ୧୭ ଆଜି ଭୁବନ ଦ୍ୱାରା ଦୁନମାନେ ଲକ୍ଷ୍ମଣପୁତ୍ରକର ଦର୍ଶନ ଲଭିବା
ଯେ ଯେ କୌଣସି ରୂପରେ ଅଭିପାଦନ। ୧୮ ଯେତେ ଶିଳ୍ପ ଦୁନମାନେ ନିଜ ଧର୍ମଦାନ ଚାକର ପାଇଁ ଅନୁପାଦ କଲିବ ଏବଂ ପ୍ରଭୁ ଯାଶୁକୁ ନିଜର ଦ୍ରାଘକ୍ରମ ରୂପେ ଗ୍ରହଣ
କରିବ, ସେତେ ଶିଳ୍ପ ସେ ଦୁନକୁ ଦର୍ଶନ ଦେବେ।'
୧୯ ପ୍ରଭୁକ ଗୁରୁକାନ୍ତନେ ନୂତନ ସୃଷ୍ଟିକାର ଚରୁପାଶୁ ପ୍ରକାଶ ଦେଲା। ଆରମ୍ଭ ହେଲା ଭୁବନ ପରାମ୍ଭ। ୨୦ ଶେଷରେ ପଢ଼ିଲା ଗୁରୁକାନ୍ତର ପାଠ। ୨୧ ଶି ବିଶାଳା ଚିନ୍ତା ଶକରର
ପୁଅ ଗୁରୁକାନ୍ତ। ନିଜ ପରିବାରର ଉଚ୍ଚା ବିଶ୍ୱସ୍ତରେ ଗୁରୁକାନ୍ତ ଶ୍ରୀକ୍ଷମ୍ପା ଗ୍ରହଣ କରୁଥିଲା। ୨୨ ଗୁରୁକାନ୍ତକୁ ପାଠି ଶିବରକୁ ପାଠୋଟି ନେଲେ ନୂତନ ପାଳକ ତେମ୍ପୁ।
୨୩ ପ୍ରଥମ ଭୁବନ ପରେ ପାଠି ଉପରକୁ ମୁଣ୍ଡ କାଟିଲା ଗୁରୁକାନ୍ତ। ୨୪ ପାଳକ ପ୍ରଶ୍ନ କଲେ- ଦେଖିଲୁ?
୨୫ ଗୁରୁକାନ୍ତର ଉତ୍ତର ଥିଲା- ନା।
୨୬ ଦୁଇଥରେ ତାକୁ ପାଠି ଶିବରେ ମାଡ଼ିବିଲେ ପାଳକ। ଆଉ ପ୍ରଶ୍ନ କଲେ- ଦେଖିଲୁ?
୨୭ ଏଥର ବି ଗୁରୁକାନ୍ତର ଉତ୍ତର ନିରୁପାଦନକ।
୨୮ ଦୁନାୟ ଥର ପାଳକ ଚିତ୍ତେ ଅଧିକ ସମୟ ପାଇଁ ଗୁରୁକାନ୍ତକୁ ପାଠି ଶିବରେ ମାଡ଼ି ବସିଲେ, ଆଉ ତା' ପରେ ଉଠାଇ ଆଣି ଦୁଇ କହିଲେ- ଏଥର ଅବଶ୍ୟ ରେଟ
ହୋଇଥିବ।
୨୯ ଦେଖି ଗୁରୁକାନ୍ତର ଉତ୍ତର ସମାନ ଥିଲା। ୩୦ ଏଣୁ ଆଉ ବିଚ୍ଚେ ଅଧିକ ସମୟ ପାଇଁ ତା' ମୁଣ୍ଡକୁ ପାଠି ଶିବରେ ମାଡ଼ିଲେ ପାଳକ। ତେବେ ବି ପ୍ରଭୁ ଯାଶୁକ ଦର୍ଶନ
ପାଇନଥିଲା ଗୁରୁକାନ୍ତ ଶବର। ୩୧ ଏହିଭଳି ଆଠ ଥର ହେଲା। ପ୍ରତିଥର ଭୁବନ ସମୟ କରୁଥାଏ। ଦେଖି ଗୁରୁକାନ୍ତର ସେଇ ଏକା ଉତ୍ତର। ୩୨ ବିଶେଷ ରପରୁ ଆଉ

Issues with the early converters

- A new converter for each encoding
- Old and incompatible for new fonts
- Unresponsive web portal
- Difficult to maintain regularly



A baby was born!

Copy text typed in Sarala font and paste below.

]cðjû^ Rûa^ ûAñ @^êZû_ Keða Gaõ_ âbê -ûgêueêê ^òRe
ZâûYK@ðû eiù_ MâjY Keða, ùiùZ gòNâ ùi Zê Kê \gð^
ù\ûaò'19_ âbêu MÊYKú@ð^ùe ^iZ^_ êAeòYúe PZê ûgßð
âKµòZ ùjfuò @úe, ùjfu Wêa^ epeúò 20 ùghue_ Wýòfû
Mêeêaúeúe_ úkòò 21 Mûñ \ògúeú WòAê gaeê_ êê
Mêeêaúeúò ^òR_ eðaúee Azú aóeêiúe Mêeêaúeú LâúÁ]cð
MâjY Keê[òfûò 22

Convert to Unicode >

Copy Unicode Odia text from the box below.

ଯୀଶୁଙ୍କ ସହ ସାକ୍ଷାତୂତ୍ୟ ବଶତଃ ପୁନଶ୍ଚ କହୁବା ଆରମ୍ଭ କଲେ । ୧୭ 'ସମସ୍ତେ
ପାପ ପଥରୁ ଫେର । କାରଣ ସ୍ୱର୍ଗରାଜ୍ୟ ନିକଟତର । ୧୭ ଆଜି ତୁବନ ଦ୍ୱାର
ତୁମ୍ଭେମାନେ ଇଶ୍ୱରପୁତ୍ରଙ୍କର ଦର୍ଶନ ଲଭିବ । ସେ ଯେ କୌଣସି ରୂପରେ

Mapping of characters

କ ଠ କ

displayed in Akruti

KUK

Corresponding ASCII
character

କ ଠ କ

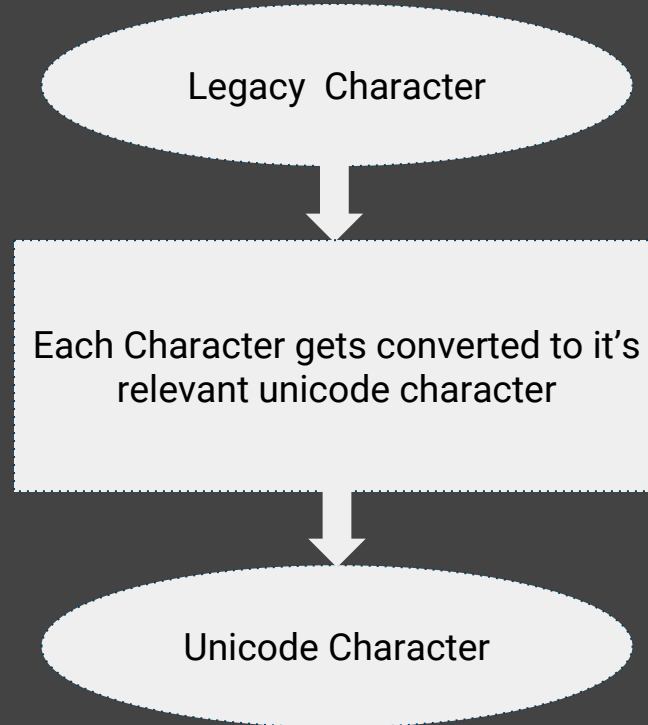
In Unicode Odia after
encoding conversion

First on-wiki converter

All converters in one page



Workflow



What we planned to change

- Responsive Design
- More Code improvement
- Modern look

Let's talk some code

- Javascript, HTML
- JQuery Introduced in new version
- Sweet Alert
- JQuery IME



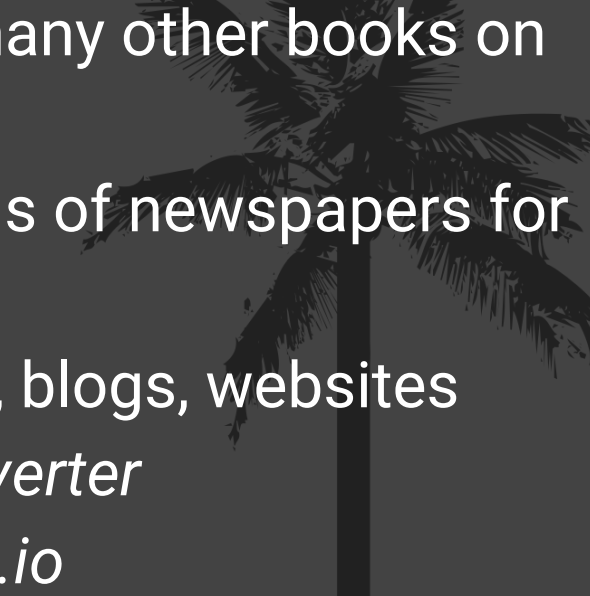
Meet our new converter

The screenshot shows the Wikipedia Odia Converter tool interface. The browser address bar displays the URL: https://or.wikipedia.org/wiki/ଉଇକିପିଡ଼ିଆ:ଚୁଲ/ଏନକୋଡ଼ିଙ୍ଗ_କନଭର୍ଟର. The page title is "ଉଇକିପିଡ଼ିଆ:ଚୁଲ/ଏନକୋଡ଼ିଙ୍ଗ କନଭର୍ଟର". The URL <https://or.wikipedia.org/s/t33> is also visible. The main content area shows a text input field with the text "ଏଠାରେ କିଛି ତାରି ଗଢ଼ନ୍ତୁ" and a "Converted Unicode text below" label. A dropdown menu is open, showing options: "Choose", "Akuti to Unicode", "GIST to Unicode", "Shreejipi to Unicode", "Apranta to Unicode", and "SambadEML to Unicode". Below the input field, there are buttons for "Akuti", "Sreejipi", "GIST", "Apranta", and "Sambad EML", along with a button that says "Click to select all and copy". A "Shortcut: WP:Converter" box is present on the right. The left sidebar contains navigation links for "ପୃଥାକ ପୃଷ୍ଠା", "ସଂପଦ ସୂଚନା ଫଳକ", "ନିର୍ଦ୍ଦେଶ ଦେବା", "ପ୍ରାକୃତକାଳୀନ ପୃଷ୍ଠା", "ଦାନ", "ପୋଷାକପୋଷାକ", "ପାଠକପଠା", "ଆଲୋଚନା ପୃଷ୍ଠା", "ସହଯୋଗ କରନ୍ତୁ", "ଉଇକି ପ୍ରକଳ୍ପ", "ପଞ୍ଜା ଛବି", "ପଞ୍ଜା ଲେଖା", "ଉଦ୍ଦେଶ୍ୟ", "ଏଠାରେ ଥିବା ଲିଙ୍କ", "ପାଠକର ଦେବା", "ଫାଇଲ ଅପଲୋଡ୍ କରିବେ", "ବିଶେଷ ପୃଷ୍ଠା", "ସହଯୋଗୀ ଲିଙ୍କ", "ପୃଷ୍ଠା ସୂଚନା", "କ୍ଷେତ୍ର URL", "ପ୍ରିଣ୍ଟ/ଉପାଳି", "କିଛି ଚିତ୍ରାଣ କରିବେ", "PDF ରାସରେ ଡାଉନଲୋଡ୍ କରିବେ", "ଛପାହୋଇପାରିବା ପୃଷ୍ଠା", "Languages", and "Add links". The bottom of the page shows the URL <https://or.wikipedia.org/wiki/WP:Converter>.

How it works



A fruit bearing tree

- Digitization of Odia Bhagabata and many other books on Odia Wikisource
 - Collecting information from e-versions of newspapers for Odia Wikipedia
 - Used by many users for social media, blogs, websites
 - *<https://or.wikipedia.org/wiki/Wp:Converter>*
 - Development: *<http://gyan111.github.io>*
- 
- A dark silhouette of a palm tree is positioned on the right side of the slide, extending from the bottom to the top. The tree has a long, thin trunk and a crown of fronds that fan out to the left and right.

Way forward

- Converting larger files
- More improvised code
- Converters for more character encoding systems
- Smart converter for auto-detection of encoding and seamless conversion





Thank You
gyana111@gmail.com

Questions?