

## NIKL Korean History Corpus

(version 1.0)

- **Title:** NIKL Korean History Corpus
- **Release Date**
  - (version 1.0) 28 June 2024
- **Data Type:** text
- **Description**
  - Content and size
    - Literature data recorded in Hangeul from the creation of Hangeul in the 15<sup>th</sup> century to the early 20<sup>th</sup> century
  - A total of 1,153
- **File format and Encoding:** XML(UTF-8 encoding)
- **Number and size of files:** 1,153 files, 149MB
- **Citation(s):**

(Korean) 국립국어원(2023). 국립국어원 국어 역사 말뭉치(버전 1.0). URL: <https://kli.korean.go.kr/corpus>

(English) The National Institute of Korean Language(2023). NIKL Korean History Corpus(v1.0)  
URL: <https://kli.korean.go.kr/corpus>
- **File Naming Rule**

order	1	2	3	4	5	6	7	8	9	10	11	12	13	14
feature	data source	type	annotation level		building year									
value	H: Historical material	X: extraction	RW: raw corpus		23: 2023									20000001 ~ 20001153 (serial number in 8 digits)

	ials				
※ Example: HXRW2320000001.xml - the first file of Korean history corpus, modified in 2023					

## • Example

<pre> &lt;doc&gt; &lt;teiHeader&gt; &lt;titleStmt&gt; &lt;title lang="kor"&gt;훈민정음&lt;/title&gt; &lt;title lang="chi"&gt;訓民正音&lt;/title&gt; &lt;volume n=""/&gt; &lt;author&gt;&lt;/author&gt; &lt;date&gt;1459&lt;/date&gt; &lt;/titleStmt&gt; &lt;notesStmt&gt; &lt;bibl&gt; &lt;title&gt;언해본 훈민정음(서강대 월인석보 권두본)&lt;/title&gt;&lt;!--국가기록유산, 디지털한글박물관 등--&gt; &lt;pubPlace&gt;서강대&lt;/pubPlace&gt; &lt;publisher&gt;&lt;/publisher&gt; &lt;date&gt;&lt;/date&gt; &lt;/bibl&gt; &lt;/notesStmt&gt; &lt;revisionDesc&gt; &lt;date&gt;&lt;/date&gt; &lt;respStmt&gt; &lt;resp&gt;입력자&lt;/resp&gt;&lt;name&gt;김송희&lt;/name&gt; &lt;/respStmt&gt; &lt;respStmt&gt; &lt;resp&gt;교정자&lt;/resp&gt;&lt;name&gt;&lt;/name&gt; &lt;/respStmt&gt; &lt;/revisionDesc&gt; &lt;/teiHeader&gt; &lt;sent type="title" lang="chi" page="1a" n="1"&gt;世宗御製訓民正音【製는 글 지을씨니 御製는 님금 지으샐 그리라】 【訓은 ㄱㄴㅇㅇㅣ오 ㅁ은 百姓ㅣ오 ㅍ은 소리니 訓民正音은 百姓 ㄱㄴㅇㅇㅣ는 正音 소리라】&lt;/sent&gt; &lt;sent type="main" lang="chi" page="1a" n="2"&gt;國之語音이【國은 나라히라 之는 입겨지라 語는 말쑤미라】 &lt;/sent&gt; &lt;sent type="main" lang="kor" page="1a" n="2"&gt;나랏 말쑤미&lt;/sent&gt; &lt;sent type="main" lang="chi" page="1b" n="1"&gt;異乎中國하야【異는 다를씨라】【乎는 아모그에 ㅎ는 겨체 ㅍ는 字 라】【中國은 皇帝 겨신 나라히니 우리 나랏 常談에 江南이라 ㅎ는 나라】&lt;/sent&gt; &lt;sent type="main" lang="kor" page="1b" n="1"&gt;中囡國꺤에 달아&lt;/sent&gt; &lt;sent type="main" lang="chi" page="1b" n="2"&gt;與文字로 不相流通홀씨【與는 이와 더와 ㅎ는 겨체 ㅍ는 字라】 【文은 글와리라】【不은 아니 ㅎ는 ㅍ디라】【相은 서르 ㅎ는 ㅍ디라】【流通은 흘러 ㅅㅁ출씨라】&lt;/sent&gt; &lt;sent type="main" lang="kor" page="1b" n="2"&gt;文字와로 서르 ㅅㅁ디 아니홀씨&lt;/sent&gt; &lt;sent type="main" lang="chi" page="2a" n="1"&gt;故로 愚民이 有所欲言하야도【故는 전치라】【愚는 어릴씨라】【有는 이실씨라】【所는 배라】【欲은 ㅎ고져 홀씨라】【言은 니를씨라】&lt;/sent&gt; &lt;sent type="main" lang="kor" page="2a" n="1"&gt;이런 전초로 어린 百꺤姓성이 니르고져 홀 배 이셔도&lt;/sent&gt; &lt;sent type="main" lang="chi" page="2b" n="1"&gt;而終不得伸其情者 多矣라【而는 입겨지라】【終은 ㅁ쑤미라】【得은 시를씨라】【伸은 펴씨라】【其는 제라】【情정은 ㅍ디라】【者는 노미라】【多는 할씨라】【矣는 말 ㅁ는 입겨지라】&lt;/sent&gt; &lt;sent type="main" lang="kor" page="2b" n="1"&gt;ㅁ쑤내 제 ㅍ들 시러 ㅍ디 ㅁ홀 노미 하니라&lt;/sent&gt; &lt;sent type="main" lang="chi" page="2b" n="2"&gt;予 爲此憫然하야【予는 내 ㅎ습시는 ㅍ디시니라】【此는 이라】【憫 然은 어엿비 너기실씨라】&lt;/sent&gt; &lt;sent type="main" lang="kor" page="2b" n="2"&gt;내 이를 爲하야 어엿비 너겨&lt;/sent&gt; &lt;sent type="main" lang="chi" page="3a" n="1"&gt;新制二十八字하노니【新은 새라】【制는 ㅁㅇ ㄱㄴㅇㅇ씨라】【二十八은 </pre>
--

스물여덟비라】</sent>  
 <sent type="main" lang="kor" page="3a" n="1">·새로 스물여덟 字를 밍 ㄱ노니</sent>  
 <sent type="main" lang="chi" page="3b" n="1">欲使人人 〇로 易習 ㅎ야 便於日用耳니라【使는 히여 ㅎ는 마리라】  
 【人은 사ㄹ미라】【易는 쉬블씨라】【習은 니길씨라】【便은 便安 ㅎ씨라】【於是 아모그에 ㅎ는 겨체 쓰는 ㄱ라】【日은 나  
 리라】【用은 ㅍ씨라】【耳는 ㅅㄹ미라 ㅎ는 ㅍ디라】</sent>  
 <sent type="main" lang="kor" page="3b" n="1">사름마다 히여 수비 니겨 날로 ㅍ매 便安키 ㅎ고저 ㅎ ㅅㄹ미  
 라</sent>  
 <sent type="main" lang="chi" page="4a" n="1">ㄱ 牙音이니 如君 ㄷ 字 初發聲 ㅎ니 竝書 ㅎ면 如蚰 ㅅ 字 初發聲  
 ㅎ니라【牙는 어미라】【如는 ㄱ 톨씨라】【初發聲은 처섬 ㅍ아 나눈 소리라】【並書는 ㄱㅅ ㅅ씨라】</sent>  
 <sent type="main" lang="kor" page="4a" n="1">ㄱ 엄쏘리니 君 ㄷ 字 처섬 ㅍ아 나눈 소리 ㄱㅅ니 ㄱㅅ ㅅ면  
 蚰 ㅅ 字 처섬 ㅍ아 나눈 소리 ㄱㅅ니라</sent>  
 <sent type="main" lang="chi" page="4a" n="2">ㅋ 牙音이니 如快 ㅊ 字 初發聲 ㅎ니라</sent>  
 <sent type="main" lang="kor" page="4a" n="2">ㅋ 엄쏘리니 快 ㅊ 字 처섬 ㅍ아 나눈 소리 ㄱㅅ니라</sent>  
 <sent type="main" lang="chi" page="4b" n="1">ㅇ 牙音이니 如業 ㅈ 字 初發聲 ㅎ니라</sent>  
 <sent type="main" lang="kor" page="4b" n="1">ㅇ 엄쏘리니 業 ㅈ 字 처섬 ㅍ아 나눈 소리 ㄱㅅ니라</sent>  
 <sent type="main" lang="chi" page="5a" n="1">ㄷ 舌音이니 如斗 ㅈ 字 初發聲 ㅎ니 竝書 ㅎ면 如草 ㅈ 字 初發聲 ㅎ  
 니라【舌은 ㅎ라】</sent>  
 <sent type="main" lang="kor" page="5a" n="1">ㄷ 혀쏘리니 斗 ㅈ 字 처섬 ㅍ아 나눈 소리 ㄱㅅ니 ㄱㅅ ㅅ면  
 草 ㅈ 字 처섬 ㅍ아 나눈 소리 ㄱㅅ니라</sent>  
 <sent type="main" lang="chi" page="5a" n="2">ㄸ 舌音이니 如吞 ㄷ 字 初發聲 ㅎ니라</sent>  
 <sent type="main" lang="kor" page="5a" n="2">ㄸ 혀쏘리니 吞 ㄷ 字 처섬 ㅍ아 나눈 소리 ㄱㅅ니라</sent>  
 <sent type="main" lang="chi" page="5b" n="1">ㄴ 舌音이니 如那 ㄴ 字 初發聲 ㅎ니라</sent>  
 <sent type="main" lang="kor" page="5b" n="1">ㄴ 혀쏘리니 那 ㄴ 字 처섬 ㅍ아 나눈 소리 ㄱㅅ니라</sent>  
 <sent type="main" lang="chi" page="6a" n="1">ㅂ 唇音이니 如擎 ㅍ 字 初發聲 ㅎ니 竝書 ㅎ면 如步 ㅍ 字 初發聲 ㅎ  
 라【唇은 입시우리라】</sent>  
 <sent type="main" lang="kor" page="6a" n="1">ㅂ 입시울쏘리니 擎 ㅍ 字 처섬 ㅍ아 나눈 소리 ㄱㅅ니 ㄱㅅ ㅅ면  
 步 ㅍ 字 처섬 ㅍ아 나눈 소리 ㄱㅅ니라</sent>  
 <sent type="main" lang="chi" page="6a" n="2">ㅍ 唇音이니 如漂 ㅍ 字 初發聲 ㅎ니라</sent>  
 <sent type="main" lang="kor" page="6a" n="2">ㅍ 입시울쏘리니 漂 ㅍ 字 처섬 ㅍ아 나눈 소리 ㄱㅅ니라  
 </sent>  
 <sent type="main" lang="chi" page="6b" n="1">ㅃ 唇音이니 如彌 ㅃ 字 初發聲 ㅎ니라</sent>  
 <sent type="main" lang="kor" page="6b" n="1">ㅃ 입시울쏘리니 彌 ㅃ 字 처섬 ㅍ아 나눈 소리 ㄱㅅ니라  
 </sent>  
 <sent type="main" lang="chi" page="7a" n="1">ㄷ 齒音이니 如卽 ㄷ 字 初發聲 ㅎ니 竝書 ㅎ면 如慈 ㄷ 字 初發聲 ㅎ  
 니라【齒는 니라】</sent>  
 <sent type="main" lang="kor" page="7a" n="1">ㄷ 니쏘리니 卽 ㄷ 字 처섬 ㅍ아 나눈 소리 ㄱㅅ니 ㄱㅅ ㅅ면 慈  
 ㄷ 字 처섬 ㅍ아 나눈 소리 ㄱㅅ니라</sent>  
 <sent type="main" lang="chi" page="7a" n="2">ㅈ 齒音이니 如侵 ㅈ 字 初發聲 ㅎ니라</sent>  
 <sent type="main" lang="kor" page="7a" n="2">ㅈ 니쏘리니 侵 ㅈ 字 처섬 ㅍ아 나눈 소리 ㄱㅅ니라</sent>