

# 『고문헌 한자 OCR 이용 매뉴얼』

2023. 05

누리IDT

## 고문헌한자시스템

본 서비스는 누리IDT 고문헌 한자 OCR 시스템 웹사이트로 고문헌 한자 자료를 보다 효율적으로 OCR/입력, 열람, 검색, 활용할 수 있도록 기획한 서비스입니다.

### 고문헌 한자 OCR 서비스

본 서비스는 고문헌 한자에 특화된 문자 인식 SI모델을 개발·적용하여 고문헌 이미지를 한자 텍스트로 추출하는 서비스입니다.

서비스 바로가기



### 한자 자형 검색 서비스

본 서비스는 한자의 음이나 부수, 획수를 알지 못하더라도 한자를 구성하는 부수와 기타 자형의 모양을 조합(선택)하여 원하는 글자를 쉽고 빠르게 찾을 수 있는 서비스입니다.

서비스 바로가기

### 누리 한자입력기

본 서비스에서 활용되는 누리한자 입력기는 한자 자형을 보고 연상되는 한글의 자음, 모음, 기호를 결합하여 간편하게 한자를 입력할 수 있는 프로그램입니다.

서비스 오픈 예정

### 이미지 검색 서비스 예시

본 서비스는 이미지를 다량 보유한 기관에서 디지털 텍스트가 없어도 OCR기술을 이용하여 이미지 속의 단어를 검색할 수 있는 서비스입니다.

※ 본 서비스는 서울대학교 규장각한국학연구원의 과학문헌 자료를 적용하였습니다.



서비스 바로가기

### OCR API 서비스 예시

본 서비스는 각 기관에서 서비스 중인 고문헌 이미지 서비스에 당사에서 제공하는 OCR API를 쉽게 추가하고 해당 이미지에서 텍스트를 추출하는 서비스입니다.



서비스 바로가기



- 고문헌 한자 OCR은 약관 동의 -> 개인정보 입력 -> 메일인증 단계를 거쳐 가입할 수 있습니다.



## 회원 가입

### 약관 동의

아래 서비스 이용약관과 개인정보 수집 및 이용에 모두 동의합니다.

### 이용약관

#### 제1장 총칙

##### 제1조 (목적)

본 이용약관(이하 '약관'이라 함)은 (주)누리아이디티(이하 '회사'라 함)가 제공하는 '누리IDT고문헌한자시스템' (https://ocr.nuriidt.co.kr) 사이트에서 디지털 콘텐츠 서비스(이하 '서비스'라 함)를 이용함에 있어 이용자(이하 '회원'이라 함)와 회사 간에 가입 및 사용에 관한 제반 사항과 기타 필요사항을 규정함을 목적으로 합니다.  
 ※ 본 약관이 적용되는 사이트는 다음과 같습니다.  
 - 고문헌 한자 OCR 서비스, 한자 자형 검색 서비스, 이미지 검색 서비스, OCR API 시범 서비스  
 (단, 제공 서비스에 따라 개별약관 동의가 필요할 수 있으며, 개별약관을 우선하여 적용합니다.)

##### 제2조 (용어의 정의)

이 약관에서 사용하는 용어의 정의는 다음과 같습니다.  
 1. '서비스'라 함은 회사가 운영하는 '누리IDT고문헌한자시스템'을 통하여 회원에게 제공되는 일체를 말합니다.

내용을 확인하였으며 동의합니다.

### 개인정보처리방침

주식회사 누리아이디티(이하 "회사")는 「개인정보 보호법」 제30조에 따라 정보주체에게 개인정보를 보호하고 이와 관련한 고충을 신속하고 원활하게 처리할 수 있도록 하기 위하여 다음과 같이 개인정보처리방침을 수립·공개합니다.

※ 이 개인정보처리방침은 2023년 3월 27일 부터 적용됩니다.

#### 제1조 개인정보 처리목적

회사는 다음의 목적을 위하여 개인정보를 처리합니다. 처리하고 있는 개인정보는 다음의 목적 이외의 용도로는 이용되지 않으며 이용 목적이 변경되는 경우에는 「개인정보 보호법」 제18조에 따라 별도의 동의를 받는 등 필요한 조치를 이행할 예정입니다.

##### 1.고객상담

- 홈페이지 "CONTACT" 메뉴를 통해 문의를 요청한 정보주체에 상담 및 문의 답변을 목적으로 개인정보를 처리합니다.
- 이메일을 통해 서비스 사용문의를 요청한 정보주체에 상담 및 문의 답변을 목적으로 개인정보를 처리합니다.

## 회원 가입

### 회원 정보 입력

\* 는 필수 입력 사항입니다.

* E-mail	<input type="text" value="로그인 및 인증에 사용할 E-mail 주소를 입력하십시오."/>	<input type="checkbox"/> 중복 확인	사용 가능한 E-mail입니다.
* E-mail 인증	<input type="text" value="2:49"/>	<input type="checkbox"/> 인증번호 발송	E-mail 인증이 필요합니다.
	인증 메일을 발송하였습니다. 인증 번호의 유효기간은 5분입니다. 인증 메일이 오지 않는 경우, 스팸에일함을 확인하세요. E-mail 인증이 완료되었습니다.	<input type="checkbox"/> 인증번호 확인	
* 성명(기관명)	<input type="text"/>		
* 비밀번호	<input type="text" value="영문과 숫자 포함 8글자 이상 입력하십시오."/>		
* 비밀번호 확인	<input type="text"/>		
* 연락처(전화번호)	<input type="text" value="-'를 포함하여 정확하게 입력하십시오."/>		
소속기관	<input type="text"/>		
정보알림 수신여부	<input type="checkbox"/> 이메일		

소중한 개인 정보(이메일, 성명, 비밀번호, 연락처)는 모두 암호화하여 저장됩니다

취소

가입신청

## 상품 구매

- 고문헌 한자 OCR은 기능 무료 테스트(30% 텍스트 추출) 사용이 가능합니다. 추가로 이용하시려면 로그인 후 원하시는 상품을 구매 후 이용하실 수 있습니다.
- 원하시는 상품을 선택한 뒤에 결재를 진행합니다.

## 상품 구매

Premium(프리미엄)을 구독하시면  
10% 할인된 가격으로 이용할 수 있습니다.

### Nuriidt OCR Lite

서비스 구독 기간 1년

텍스트 추출 횟수 500회

한자 자형 검색 서비스 무제한

---

**500,000원**  
(VAT 별도)

선택

### Nuriidt OCR Premium

10% 할인!

서비스 구독 기간 1년

텍스트 추출 횟수 5,500회

한자 자형 검색 서비스 무제한

---

**5,000,000원**  
(VAT 별도)

선택

# 1. 고문헌 한자 OCR 서비스\_메인 화면

누리DT 고문헌한자시스템

고문헌 한자 OCR Tip 도움말

**1 원문 이미지**



GK12444\_00\_0\_0001\_008a.jpg 세크런트 표시

이미지 파일 선택

**2 원문 텍스트**

安城君李叔蕃自成大功後驕傲視同列幸權不啻奴隸君命召則稱疾不往中使候者絡繹不絕而絲竹聞於內室或欲注人官爵則書名於小簡伴人奏之以故親朋布列疆仕大起甲第於敦義門內惡人馬之聲奏塞門禁行人奢侈日甚遂得罪杖流咸陽別墅世宗命儒臣撰龍飛御天歌以叔蕃知太宗朝事馳駘召之叔蕃以白衣詣闕達官權宰皆後生爭袪拜謁叔蕃但揮手止之日少時某也英邁某也信實余亦知其為令長之器果然其善氣榮榮曾不少屈

獻陵廟世子讓寧時引入大臣以其事語之黃喜李復時為判書堅執不可左從俄謫千外凡六年柳廷顯獨贊其議英廟立首入為相獻陵一日命召黃喜至戴楯高頂笠子穿碧色布圍領帶藍綠兒詣承政院喜方自田野而至形貌儼然而已入未甚奇五

텍스트 파일로 내려받기

**3 한자정보 및 수정**

馬	100%	적용하기
말, 벼슬 이름, 아자말어		
유니코드	U+099AC	
한자	馬	
薦	0.00%	적용하기
죽 두릅다		
유니코드	U+04500	
표준	薦	
國	0.00%	적용하기
국 나라		
유니코드	U+056EF	
표준	國	
駟	0.00%	적용하기
어 말 부리다, 말어거		
유니코드	U+099AD	
한자	駟	
駒	0.00%	적용하기
구 말아지		
유니코드	U+09902	
한자	駒	

이 사이트는 크롬 브라우저에 최적화되어 있습니다. 다만, 폰트 지원 문제로 특정 한자 영역의 일부 글자가 깨져 보이는 경우가 있습니다.

## 1. 이미지 업로드 창

- 추출하고자 하는 이미지를 업로드합니다.
- 탐색기에서 이미지를 "드래그-드롭"해도 됩니다.

## 2. 원문 텍스트 창

- 1번 창에 업로드한 이미지를 AI가 인식하여 추출한 한자입니다.

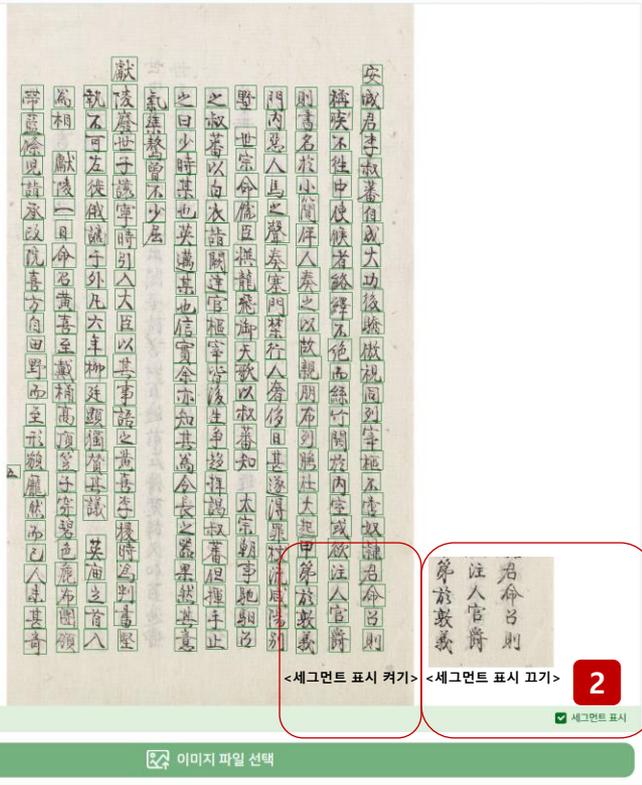
## 3. 한자 정보 및 수정 창

- 추출된 한자의 정보가 나타납니다.
- AI가 추천한 자형 5글자를 보여줍니다.
- 추출된 한자는 수정할 수 있습니다.

# 1-1. 고문헌 한자 OCR 서비스\_원문 이미지 업로드

원문 이미지

## 1 이미지 예시



※ 규장각원문검색서비스\_소대수언

## 고문헌 이미지 업로드

1. 파일 형식 : JPG, PNG
2. 파일 용량 : 20Mb 이하
  - 20Mb 이상 시, 속도가 저하될 수 있습니다.
3. 고문헌 이미지 서체 : 해 · 행서
  - 해서와 행서체의 한자를 주로 학습하여 개발되었기 때문에 초서체나 다른 서체는 인식하지 못하거나 인식률이 낮을 수 있습니다.
4. **2** 에서 세그먼트 표시의 '켜기/끄기'가 가능합니다.
5. 업로드 시, 1번 예시와 같이 단면만을 업로드합니다.
  - 양면 업로드 시, 인식률이 저하될 가능성이 있습니다.
6. 이미지가 왜곡되지 않고, 밝고 선명할수록 인식률이 좋습니다.
  - 기울기 조정, 밝기 조정 등

# 1-2. 고문헌 한자 OCR 서비스\_원문 텍스트 추출

원문 이미지

원문 텍스트

1 이미지 위의 박스(세그먼트)에 마우스 포인터가 있으면 원문 텍스트에도 박스가 표시됩니다. 이때 인공지능이 추천한 글자가 바로 위에 보입니다.

2 원문 텍스트에 마우스 포인터가 있으면 이미지 위의 박스(세그먼트)에도 박스가 표시됩니다. 이때 인공지능이 추천한 글자 5개가 확률이 높은 순으로 나타납니다.

3 클릭하면 이미지와 원문 텍스트의 해당 위치가 반전되어 표시됩니다

4 정자로 변환하기 기능은 이체자를 모두 정자로 변환해 줍니다.

5 텍스트 파일로 내려받기 기능으로 원문 텍스트를 파일로 저장할 수 있습니다.

- 1 이미지 위의 박스(세그먼트)에 마우스 포인터가 있으면 원문 텍스트에도 박스가 표시됩니다. 이때 인공지능이 추천한 글자가 바로 위에 보입니다.
- 2 원문 텍스트에 마우스 포인터가 있으면 이미지 위의 박스(세그먼트)에도 박스가 표시됩니다. 이때 인공지능이 추천한 글자 5개가 확률이 높은 순으로 나타납니다.
- 3 클릭하면 이미지와 원문 텍스트의 해당 위치가 반전되어 표시됩니다
- 4 정자로 변환하기 기능은 이체자를 모두 정자로 변환해 줍니다.
- 5 텍스트 파일로 내려받기 기능으로 원문 텍스트를 파일로 저장할 수 있습니다.

※ 규장각원문검색서비스\_소대수연

다운로드

# 1-3. 고문헌 한자 OCR 서비스\_한자 정보

## 원문 텍스트

## 한자정보 및 수정

安城君李叔蕃自成大功後驕傲視同列宰樞不啻奴 **2** 命召  
 則  
 稱疾不往中使候者絡繹不絕而絲竹關於內室或欲注人官爵  
 則書名於小簡伴人奏之以故親朋布列膳仕大起甲 **3** 敦義  
 門內惡人馬之聲奏塞門禁行入奢侈日甚遂得罪杖流咸陽別  
 墅世宗命儒臣撰龍飛御天歌以叔蕃知太宗朝事馳駟召  
 之叔蕃以白衣詣闕達官樞宰皆後生爭趨拜謁叔蕃 **4** 手止  
 之曰少時某也英運某也信實余亦知其爲令長之器果然其意  
 氣桀驁曾不少屈  
**1**  
 獻陵廢世子讓寧時引入大臣以其事語之黃喜李稷 **5** 判書  
 堅  
 執不可左徙俄謫 **千** 外凡六年柳廷顯獨贊其議英廟立首入  
 爲相獻陵 **千于乎手子** 高頂笠子穿碧色氈布團領  
 帶藍條兒 **千于乎手子** 而至形貌龐然而已入未甚奇  
 五

正字로 변환하기    텍스트 파일로 내려받기

## 한자정보

- 1** 인공지능이 추천한 글자 5개가 확률이 높은 순으로 나타납니다.
- 2** 뜻과 음
- 3** 유니코드
- 4** 정자, 이체자, 간자, 동자, 고자 등
- 5** SI의 추천 확률
- 6** 추천 글자에 없는 글자로 수정할 때

# 1-4. 고문헌 한자 OCR 서비스\_수정 기능\_텍스트 수정 기능

**원문 이미지**

**원문 텍스트**

安城君李叔蕃自成大功後驕傲同列幸權不啻奴隸君命召別稱疾不往中使候會結譯不絕而絲竹關於內室或欲注人賞罰則害名於小簡伴人美之以故戰戰兢兢不敢言事於殿門內惡人馬之聖美聖門禁行人有修日甚連連罪杖流威隔別墅世宗命臣儒能飛御天歌以祝蕃知大宗廟事馳船召之叔蕃以白衣殿前連官權幸後生爭移祥龍絙蕃但揮手止之日少時某也英運某也價賈余亦知其為令長之淵果然其意氣兼兼曾不少屈

獻陵廢世子諱寧時引入大臣以其事語之黃喜李穰時為判書監執不可左從俄謫于外凡六年柳瑤稱贊其諱英山立首入為相獻陵一日命召黃喜至龍榻高頂笠子穿鞞色魚布兩領帶藍纒兒詣承政院喜方自田對而至形影蕭然而已入未甚奇五

**한자정보 및 수정**

千	80.10%	적용하기
千	일한	
유니코드	U+05343	
于	19.63%	적용하기
우	여조사	
유니코드	U+04E8E	
眞體字	ㄅ ㄅ ㄅ	
乎	0.24%	적용하기
호	여조사	
유니코드	U+04F4E	
眞體字	ㅎ	
手	0.03%	적용하기
수	손, 잡다, 가지다, 사방	
유니코드	U+0624B	
眞體字	루 ㅁ ㅁ ㅁ	
子	0.00%	적용하기
자	아들, 사식, 용짓말	
유니코드	U+05B50	
眞體字		
작성 입력		적용하기

## 수정 기능

### 1. 텍스트 수정 기능

○ 잘못 인식된 한자가 추출 시, 수정 가능합니다.

- 1 - AI 추천 자형에 한자가 있을 시
- 2 - AI 추천 자형에 한자가 없을 시

## 텍스트 수정 기능 예시



# 1-5. 고문헌 한자 OCR 서비스\_수정 기능\_정자로 변환 기능

## 수정 기능

### 2. 정자로 변환 기능

원문 텍스트

원문 텍스트

安城君李叔善自成大功後驕傲視同列幸權不啻奴隸君命召則稱疾不往中使僕者絡繹不絕而絲竹關於內室或欲注人宮爵則書名於小簡伴人奏之以故親朋布列驢仕大起甲第於敦義門內惡人馬之聲奏塞門禁行人者修日甚遠得罪杖流咸陽別墅世宗命僕臣撰能飛御天歌以叔善知太宗朝事馳驅召之叔善以白衣詣開達官權幸皆後生爭趨拜謁叔善但揮手止之日少時某也英運某也價賈余亦知其為令長之器果然其意氣榮驚曾不少屈

獻陵廢世子讓寧時引入大臣以其事語之美宮李穆時為判書堅執不可左徒俄請外凡六年柳廷顯携贊其語英命立首入為相獻陵一日命召黃喜至戴桶高頂笠于穿碧色羅布圍領帶藍絛兒詣承政院宮方自田野而至形貌龐然而已入未甚奇五

정자로 변환하기    텍스트 파일로 내려받기

정자 변환 전

蕃知太宗朝事  
生争趨拜謁叔  
知其為令長之

원문 텍스트

원문 텍스트

安城君李叔善自成大功後驕傲視同列幸權不啻奴隸君命召則稱疾不往中使僕者絡繹不絕而絲竹關於內室或欲注人宮爵則書名於小簡伴人奏之以故親朋布列驢仕大起甲第於敦義門內惡人馬之聲奏塞門禁行人者修日甚遠得罪杖流咸陽別墅世宗命僕臣撰能飛御天歌以叔善知太宗朝事馳驅召之叔善以白衣詣開達官權幸皆後生爭趨拜謁叔善但揮手止之日少時某也英運某也價賈余亦知其為令長之器果然其意氣榮驚曾不少屈

獻陵廢世子讓寧時引入大臣以其事語之美宮李穆時為判書堅執不可左徒俄請外凡六年柳廷顯携贊其語英命立首入為相獻陵一日命召黃喜至戴桶高頂笠于穿碧色羅布圍領帶藍絛兒詣承政院宮方自田野而至形貌龐然而已入未甚奇五

정자로 변환하기    텍스트 파일로 내려받기

정자 변환 후

蕃知太宗朝事  
生爭趨拜謁叔  
知其為令長之

정자 변환

### 정자 변환

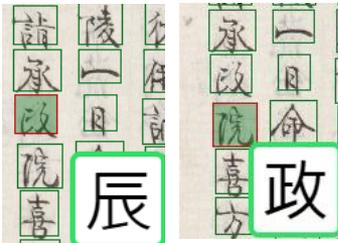
- 인식된 한자는 원문 이미지의 자형을 그대로 반영합니다.
- 원문 텍스트 하단 '정자로 변환하기' 버튼을 클릭하면 정자로 변환이 가능합니다.
- 모든 이체자는 정자로 변환되며, 우측 이미지와 같이 빨간색으로 변환된 한자가 나타납니다.

### ※ 주의점 예시

兒詣承政院    兒詣承辰政院

수정 기능 '직접 입력'을 통해 '承+辰'을 추가하여 수정

정자 변환



원문 이미지의 세그먼트와 인식된 한자 순서 어긋남.

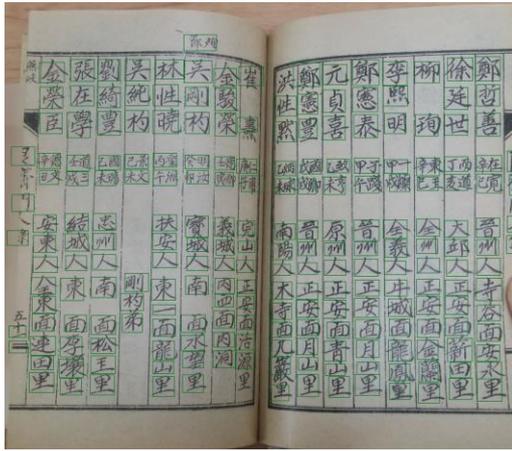
### ※ 주의점

- 오인식된 한자를 수정하기 전에 반드시 수정 전에 정자로 변환해야 합니다.
- 이체자로 수정 후에 정자 변환 시, 이체자는 정자로 변환됩니다.
- 한 글자를 여러 개의 글자로 수정 후 정자 변환 원문 이미지의 세그먼트와 인식된 한자의 순서가 어긋납니다.
- 수정이 다 끝난 후 새로고침(F5)을 누를 시, 처음 인식했던 그대로 돌아옵니다.

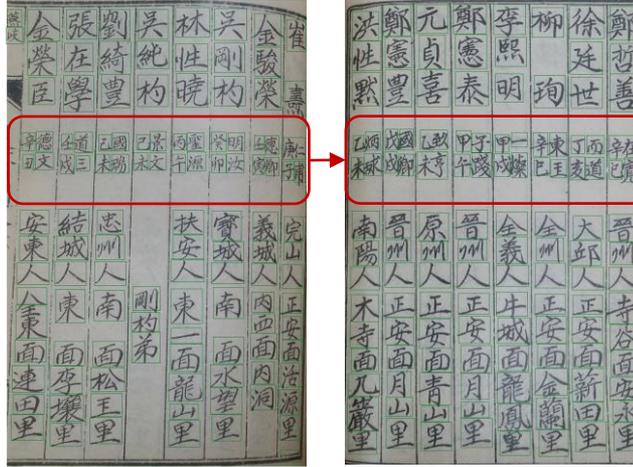
- 이용자가 업로드한 고문헌 이미지 내에서 인식하지 못하거나 다르게 인식하는 한자가 있을 수 있습니다.
- 일반적인 고도서 형태인 세로쓰기로 제작된 고문헌 이미지의 인식에 최적화되어 있습니다.
- 고문헌 이미지가 양면인 경우 2면으로 분할하여 업로드 하면 양면보다 문자 인식률이 더 좋아집니다.
- 가로쓰기나 세로쓰기 10줄 이상의 형태로 제작된 고문헌은 인식하지 못하거나 인식률이 낮을 수 있습니다.
- ‘해서체’와 ‘행서체’ 한자를 주로 학습하여 개발되었기 때문에 초서체나 기타 다른 서체(전서체, 예서체 등)로 제작된 고도서는 인식하지 못하거나 인식률이 낮을 수 있습니다.
- 금석문 등(음각이나 반전(검은 바탕에 흰 글자))으로 제작된 고문헌 이미지는 인식하지 못하거나 인식률이 급격히 떨어질 수 있습니다.
- 이용 가능한 이미지 파일 포맷은 jpg, png 입니다.(20Mb 이상의 이미지 파일은 속도가 저하될 수 있습니다.)
- 고문헌 이미지가 왜곡(기울어짐, 돌아감 등)되지 않고, 밝고 선명할수록 인식률이 좋습니다.(기울기를 조정)
- 고문헌 이미지 바탕이 어두우면 바탕색을 밝게 조정하면 인식률에 도움이 됩니다.(해상도 조정)
- 다운로드한 파일을 메모장에서 열면 화면의 글꼴 때문에 확장 한자 영역의 일부 글자가 깨져 보이는 경우가 있습니다.

# [주의] 인식을 저하 예시 이미지

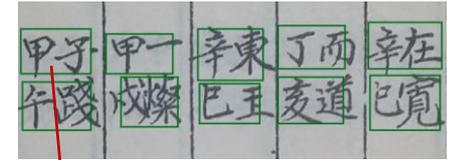
※ 양면 스캔 이미지는 인식이 낮을 수 있습니다.



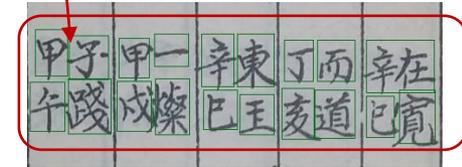
양면을  
분할하여  
업로드



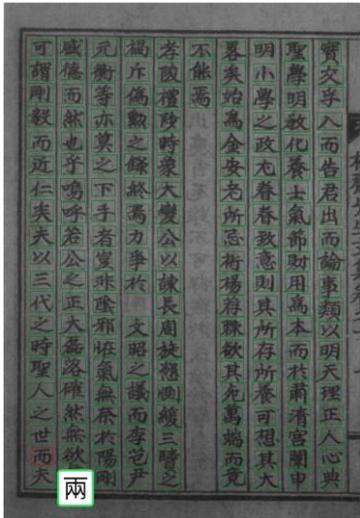
단면으로 분할하여 업로드하면 인식이 좋아집니다.



세그먼트가 잘 됨



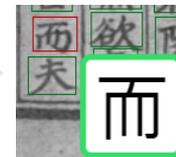
※ 이미지가 어둡거나 흐리면 인식이 낮습니다.



명도를  
조절하여  
업로드



인식을  
상승



※ 인식을 좋게 하려면

- 양면을 하나로 스캔한 이미지는 각각 스캔하여 업로드
- 어두운 이미지는 밝기와 채도를 조절하여 업로드
- 기울어진 이미지는 수직으로 조절하여 업로드

## 고문헌한자시스템

본 서비스는 누리IDT 고문헌 한자 OCR 시스템 웹사이트로 고문헌 한자 자료를 보다 효율적으로 OCR/입력, 열람, 검색, 활용할 수 있도록 기획한 서비스입니다.

### 고문헌 한자 OCR 서비스

본 서비스는 고문헌 한자에 특화된 문자 인식 SI모델을 개발·적용하여 고문헌 이미지를 한자 텍스트로 추출하는 서비스입니다.

서비스 바로가기



### 한자 자형 검색 서비스

본 서비스는 한자의 음이나 부수, 획수를 알지 못하더라도 한자를 구성하는 부수와 기타 자형의 모양을 조합(선택)하여 원하는 글자를 쉽고 빠르게 찾을 수 있는 서비스입니다.

서비스 바로가기

### 누리 한자입력기

본 서비스에서 활용되는 누리한자 입력기는 한자 자형을 보고 연상되는 한글의 자음, 모음, 기호를 결합하여 간편하게 한자를 입력할 수 있는 프로그램입니다.

서비스 오픈 예정

### 이미지 검색 서비스 예시

본 서비스는 이미지를 다량 보유한 기관에서 디지털 텍스트가 없어도 OCR기술을 이용하여 이미지 속의 단어를 검색할 수 있는 서비스입니다.

※ 본 서비스는 서울대학교 규장각한국학연구원의 과학문헌 자료를 적용하였습니다.



서비스 바로가기

### OCR API 서비스 예시

본 서비스는 각 기관에서 서비스 중인 고문헌 이미지 서비스에 당사에서 제공하는 OCR API를 쉽게 추가하고 해당 이미지에서 텍스트를 추출하는 서비스입니다.



서비스 바로가기

## 2. 한자 자형 검색 서비스 안내

### 이런 경우에 이용해보세요!

- 한문 고문헌을 읽다가 음이나 뜻을 모르는 한자가 있을 때
- 정자와 이체자, 속자 등의 정보를 알고 싶을 때
- 한자의 음이나 부수, 획수를 알지 못해 한자를 찾기 어려울 때
- 한자 자형 검색 서비스에게 맡겨보세요.

누리IDT 고문헌한자시스템

한자 자형 검색

자형 목록 | 자형 상세보기 | 자형 모아보기

한자 상세보기

한자정보

韓

韋部 8획 (총 17획)

한 유물·난간, 중국에 있던 나라 이름, 우리나라 이름

유니코드 U+097D3

한자 韓

- 한자의 음이나 부수, 획수를 알지 못하더라도 한자 자형의 조합으로 검색할 수 있는 서비스입니다.



## 2-2. 한자 자형 검색

자형 목록    자형 펼쳐보기    자형 모아보기

이제 자형을 선택하면, 관련된 자형이 자동으로 보여집니다.

한자 상세보기    한자정보

### 자형 검색

한자의 구성 형태를 목록으로 제공한 부수를 보여줍니다.

글자의 자형을 조합(선택)하여 찾고자 하는 글자를 찾습니다.

- 자형 펼쳐보기 : 한글 자 모음과 비슷한 글자로 나열된 자형
- 자형 모아보기 : 펼쳐보기의 자형이 한눈에 보이도록 나열됨
- 한자 상세 보기 : 조합에 맞는 한자들이 모두 보입니다.
  - 조합된 한자는 획순으로 나열
- 한자 정보 : 한자 1글자의 정보 훈과 음, 이체자 및 유니코드

자형 펼쳐보기    자형 모아보기

자형 목록    자형 펼쳐보기    자형 모아보기

이제 자형을 선택하면, 관련된 자형이 자동으로 보여집니다.

한자 상세보기    한자정보

한자 상세보기    총 78건

4획

天

6획

关

7획

吳 呑

8획

泰 昊 悉

9획

朕 朕 朕 奕 侯

10획

朕 啞 泰 沃 娛 糕

11획

迭 悞 焔 朕 啞 詠 括

12획

天 吳 泰 昊 悉 朕 朕 朕 奕 侯 朕 啞 泰 沃 娛 糕 迭 悞 焔 朕 啞 詠 括

### 한자정보

关

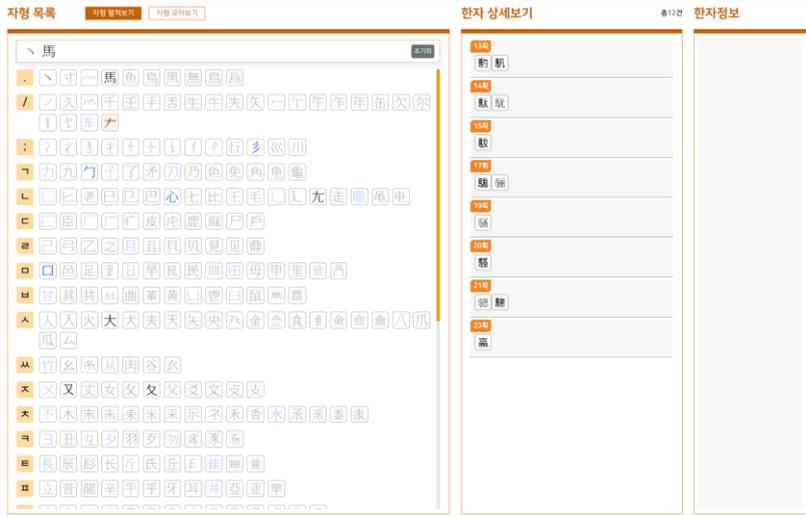
大部 3획 (총 6획)

관 문 빛장, 목, 중요하다, 관계하다

유니코드 U+05173

正字 關

# 2-3. 한자 자형 검색\_예시



## 자형 선택

- 찾고자 하는 한자의 자형을 선택합니다.
  - **1** 과 같이 자형 선택 시, 각 한자의 조합에 관련 없는 자형은 회색으로 변경되며, 선택할 수 없습니다.
  - 자형 조합은 순서에 상관없이 검색이 가능합니다.
  - **2** 과 같이 선택된 자형을 같은 자형으로 한번 더 선택 시, 조합에서 삭제되어 다른 자형을 선택하여 검색할 수 있습니다.

## 자형 검색 예시.

**예시) 大韓民國**

人 人 入 火 大 犬 夫 天 矢 央 夂 金 金 食 倉

大 大

白 自 血 由 鬼 卑 身 舟 良 阜 鼻 鹵 言 車

土 土 鼓 艸 去 赤 无 韋 韦 虫 一 方 六 亡

車 + 韋 = 韓

口 邑 足 趾 日 早 良 民 血 田 母 甲 里 骨 尸

民 民

口 邑 足 趾 日 早 良 民 血 田 母 甲 里 骨 尸

卜 片 耳 耳 阝 戈 气 飞 飛

口 + 戈 = 國

**예시) 李舜臣**

木 不 木 未 未 未 米 采 示 禾 禾 香 水 丞 丞

力 九 子 了 矛 刀 乃 色 兔 角 龟 龜

木 + 子 = 李

久 六 千 壬 手 舌 生 牛 失 矢 尸 尔 午

一 一 二 三 豆 鬲 至 云

丷 + 一 = 舜

匚 臣 厂 广 疒 皮 疒 鹿 麻 尸 尸

臣 = 臣

## 고문헌한자시스템

본 서비스는 누리IDT 고문헌 한자 OCR 시스템 웹사이트로 고문헌 한자 자료를 보다 효율적으로 OCR/입력, 열람, 검색, 활용할 수 있도록 기획한 서비스입니다.

### 고문헌 한자 OCR 서비스

본 서비스는 고문헌 한자에 특화된 문자 인식 SI모델을 개발·적용하여 고문헌 이미지를 한자 텍스트로 추출하는 서비스입니다.

서비스 바로가기



### 한자 자형 검색 서비스

본 서비스는 한자의 음이나 부수, 획수를 알지 못하더라도 한자를 구성하는 부수와 기타 자형의 모양을 조합(선택)하여 원하는 글자를 쉽고 빠르게 찾을 수 있는 서비스입니다.

서비스 바로가기

### 누리 한자입력기

본 서비스에서 활용되는 누리한자 입력기는 한자 자형을 보고 연상되는 한글의 자음, 모음, 기호를 결합하여 간편하게 한자를 입력할 수 있는 프로그램입니다.

서비스 오픈 예정

### 이미지 검색 서비스 예시

본 서비스는 이미지를 다량 보유한 기관에서 디지털 텍스트가 없어도 OCR기술을 이용하여 이미지 속의 단어를 검색할 수 있는 서비스입니다.

※ 본 서비스는 서울대학교 규장각한국학연구원의 과학문헌 자료를 적용하였습니다.



서비스 바로가기

### OCR API 서비스 예시

본 서비스는 각 기관에서 서비스 중인 고문헌 이미지 서비스에 당사에서 제공하는 OCR API를 쉽게 추가하고 해당 이미지에서 텍스트를 추출하는 서비스입니다.



서비스 바로가기

# 3. 이미지 검색 서비스 안내

## 이미지 DB의 한자를 검색해보세요!

- 한자(고문헌) 원문 이미지를 다량 보유한 기관에 적용되는 서비스로서, OCR기술을 이용하여 원문 텍스트 정보 추출 후 글자 정보 및 글자의 위치값을 포함한 DB를 생성합니다. 생성된 DB를 검색하여 해당 검색어가 있는 원문 이미지를 목록으로 제공하고, 텍스트 결과를 제공합니다.



**검색 결과**

원문이미지 검색 결과 총 5페이지가 검색되었습니다.

과학문헌 이미지 검색시범 서비스

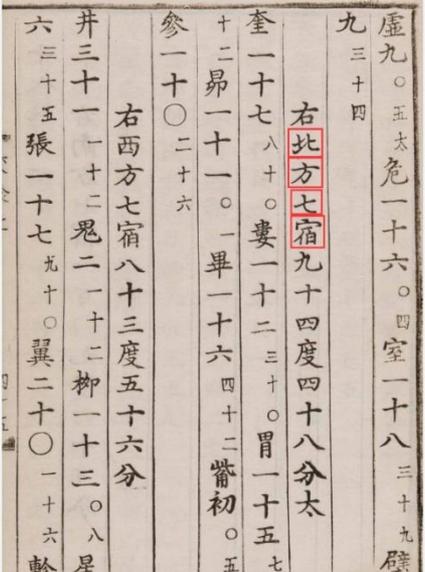
검색어:

원문 이미지 검색 서비스는 한자(漢字) 검색만 가능합니다.



<p><b>文食推歩法  교식추보법</b></p> <p>원문: ...五太危一十六四室一十八三十九星九十三右北方七宿九十四度四十八分大星一十七八十星一十二三十星一十五七十一星一十一星一十六四十二星初五星一二十六右西方七宿八十三度...</p>	<p>이미지 보기</p>
<p><b>重修大明曆 중수대명력</b></p> <p>원문: ...度少六十八危十五度半星七度星八度北方七宿九十四度六十八星十六度半星十二度星計五度星十一度星十七度少兩半星十度半右西方七宿八十三度并三針少二度半...</p>	<p>이미지 보기</p>
<p><b>重修大明曆 중수대명력</b></p> <p>원문: ...度少六十八危計六度星八度少星九度半右北方七宿九十四度六十八星計七度大星十二度大星十五度半星十一度星計六度半星參星大右西方七宿八十三度大一百七十七五六八...</p>	<p>이미지 보기</p>
<p><b>庚午元曆 경오원력</b></p> <p>원문: ...少度九少六十七危十五半星七度星八度北方七宿九十四度六十七星計十六度星十二度十五星一十一度半星計十半右西方七宿八十三度并三十三少星二半星十三大星六十七...</p>	<p>이미지 보기</p>
<p><b>庚午元曆 경오원력</b></p> <p>원문: ...十一度九少六十七危十六星八少星九半右北方七宿九十四度六十七星計十七大星十二大星十五半星一十一度半星計十半右西方七宿八十三度大并三十三半星二半星十三少星六十七...</p>	<p>이미지 보기</p>

**결과 확인**



右西方七宿八十三度五十六分

井三十一 一十二 鬼二十一 十二 柳一十三 〇 入星

六三十五 張一十七 九十 〇 翼二十一 〇 一十六 軫

參一十 〇 二十六

奎一十七 八 十 〇 婁一十二 三 十 〇 胃一十五 七

十二 昴一十一 〇 一畢一十六 四 十二 觜初 〇 五

右北方七宿九十四度四十八分太

九 三 十 四 〇 五 太 危 一 十 六 〇 四 室 一 十 八 三 十 九 壁

- 이미지를 컴퓨터에서 처리할 수 있는 텍스트로 변환하는 서비스입니다.
- 인공지능 OCR을 이용하여 텍스트가 없어도 이미지 DB의 한자를 검색할 수 있는 서비스입니다.

# 3-1. 이미지 검색\_예시

- 인공지능 OCR을 이용하여 텍스트가 없어도 이미지 DB의 한자(단어, 문장)를 검색할 수 있습니다.
- OCR로 인식된 한자의 위치가 **붉은글자**로 표시됩니다.
- 검색된 한자(단어, 문장)의 **원문 문헌 정보**를 알 수 있습니다.



과학문헌 이미지 검색시범 서비스

北方七宿

※ 과학문헌 이미지 검색 시범 서비스는 한자(漢字) 검색만 가능합니다

원문이미지 검색 결과 총 5페이지가 검색되었습니다.

**交食推步法 교식추보법** 이미지 보기

청구기호 : 奎11241 편저자 : 李純之·金石煥(朝鮮) 受命編 / 이순지·김석재(조선) 수명편 간행년대: 1458 사부분류 : 子部 天文·算法類 天文

**OCR text** ...五太危一十六四室一十八三十九壁九三十四右北方七宿九十四度四十八分太奎一十七八...

**重修大明曆 중수대명력** 이미지 보기

청구기호 : 奎12441 편저자 : [李純之(朝鮮)·金淡(朝鮮) 受命校正] / [이순지(조선)·김담(조선) 수명교정] 간행년대: 1418 사부분류 : 子部 天文·算法類 天文

**OCR text** ...度少秋六十八危十五度半室十七度壁八度太右北方七宿九十四度秘六十八室十六度半婁十...

**重修大明曆 중수대명력** 이미지 보기

청구기호 : 奎12441 편저자 : [李純之(朝鮮)·金淡(朝鮮) 受命校正] / [이순지(조선)·김담(조선) 수명교정] 간행년대: 1418 사부분류 : 子部 天文·算法類 天文

**OCR text** ...度少秘六十八危計六度室卜八度少壁九度半右北方七宿九十四度六十八秘室計七度太婁十...

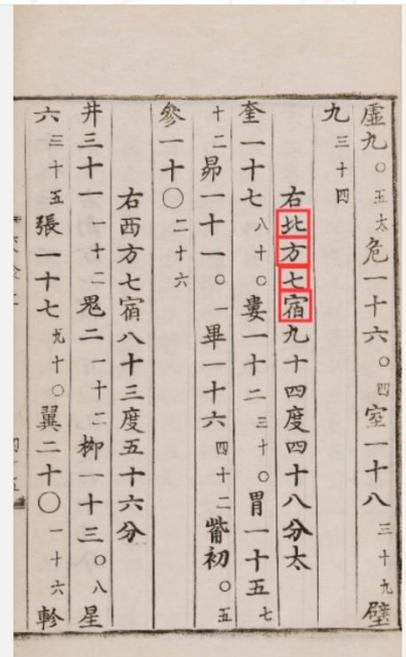
**庚午元曆 경오원력** 이미지 보기

청구기호 : 奎12443 편저자 : 耶律楚材(元) 編,李純之(朝鮮)·金淡(朝鮮) 受命校正 / 아음초재(원) 편,이순지(조선)·김담(조선) 수명교정 간행년대: 1418 사부...

**OCR text** ...一少虛九少六十七秘危十五半室十七壁八太右北方七宿九十四度六十七秘室十六半婁十二...

**庚午元曆 경오원력** 이미지 보기

검색 예시 ▶  
- 北方七宿 검색



## 고문헌한자시스템

본 서비스는 누리IDT 고문헌 한자 OCR 시스템 웹사이트로 고문헌 한자 자료를 보다 효율적으로 OCR/입력, 열람, 검색, 활용할 수 있도록 기획한 서비스입니다.

### 고문헌 한자 OCR 서비스

본 서비스는 고문헌 한자에 특화된 문자 인식 SI모델을 개발·적용하여 고문헌 이미지를 한자 텍스트로 추출하는 서비스입니다.

서비스 바로가기



### 한자 자형 검색 서비스

본 서비스는 한자의 음이나 부수, 획수를 알지 못하더라도 한자를 구성하는 부수와 기타 자형의 모양을 조합(선택)하여 원하는 글자를 쉽고 빠르게 찾을 수 있는 서비스입니다.

서비스 바로가기

### 누리 한자입력기

본 서비스에서 활용되는 누리한자 입력기는 한자 자형을 보고 연상되는 한글의 자음, 모음, 기호를 결합하여 간편하게 한자를 입력할 수 있는 프로그램입니다.

서비스 오픈 예정

### 이미지 검색 서비스 예시

본 서비스는 이미지를 다량 보유한 기관에서 디지털 텍스트가 없어도 OCR기술을 이용하여 이미지 속의 단어를 검색할 수 있는 서비스입니다.

※ 본 서비스는 서울대학교 규장각한국학연구원의 과학문헌 자료를 적용하였습니다.



서비스 바로가기

### OCR API 서비스 예시

본 서비스는 각 기관에서 서비스 중인 고문헌 이미지 서비스에 당사에서 제공하는 OCR API를 쉽게 추가하고 해당 이미지에서 텍스트를 추출하는 서비스입니다.



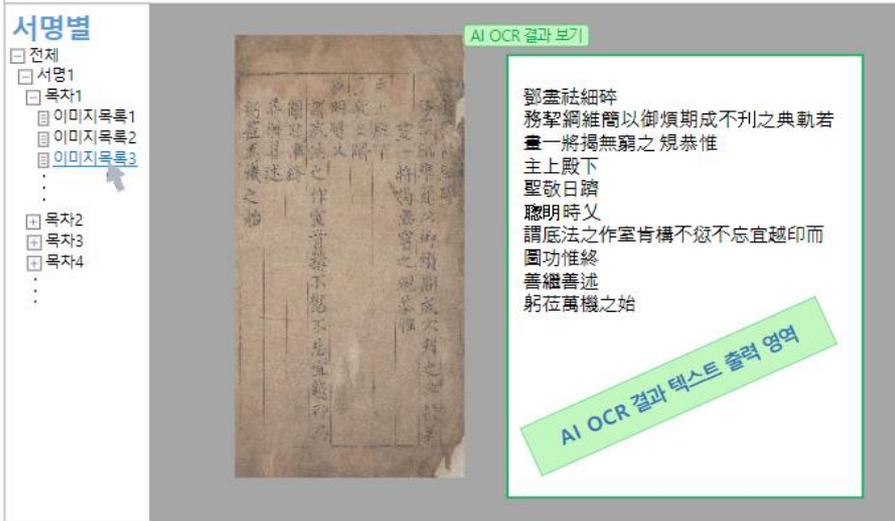
서비스 바로가기

# 4. OCR API

## 1. 고객기관 자체 원문 이미지 뷰어에서 이용하는 경우

- 고객기관은 고객기관의 자체 원문 이미지 뷰어가 있어야 합니다.
- '결과 출력 영역'를 추가 개발
  - 고객기관의 원문 이미지 뷰어에 텍스트 결과가 출력됩니다.
- 고객기관에서는 **AI OCR 결과 보기 버튼 추가**
  - 고객 기관에서 삽입(Javascript 등 이용)하고 API와 연동해야 하며, API 연동을 위한 프로그램 코드는 당사에서 제공합니다.

헤더 영역(CI, BI 및 뷰어기능버튼 등이 위치)



## 2. OCR 결과 전용 뷰어를 이용하는 경우

- 고객기관은 고객기관의 자체 원문 이미지 뷰어가 있어야 합니다.
- 이미지 뷰어에 **AI OCR 결과보기 버튼만을 추가**
  - 고객 기관에서 삽입(Javascript 등 이용)하고 API와 연동해야 하며, API 연동을 위한 프로그램 코드는 당사에서 제공합니다.
  - 누리 IDT의 뷰어를 불러와 결과가 출력됩니다.





## 이용 문의 안내

고문헌한자시스템 이용과 관련하여 다른 궁금하신 점이 있나요?  
궁금하신 점은 아래의 연락처로 언제든지 문의하여 주세요.  
담당자가 신속하고 정확하게 확인하여 답변드리겠습니다.

### ❖ 누리 IDT 문의 이용 안내

- 전화 : 070-7405-3332
- E-mail : [nuriidt@nuriidt.co.kr](mailto:nuriidt@nuriidt.co.kr)