

# Chinese Characters Dict link Telegram Bot

COSCUP 2016 - Lightning Talk

성대현(DaeHyun Sung/成大鉉/成大鉉/ソンデヒョン)

# Sponsor of CJK



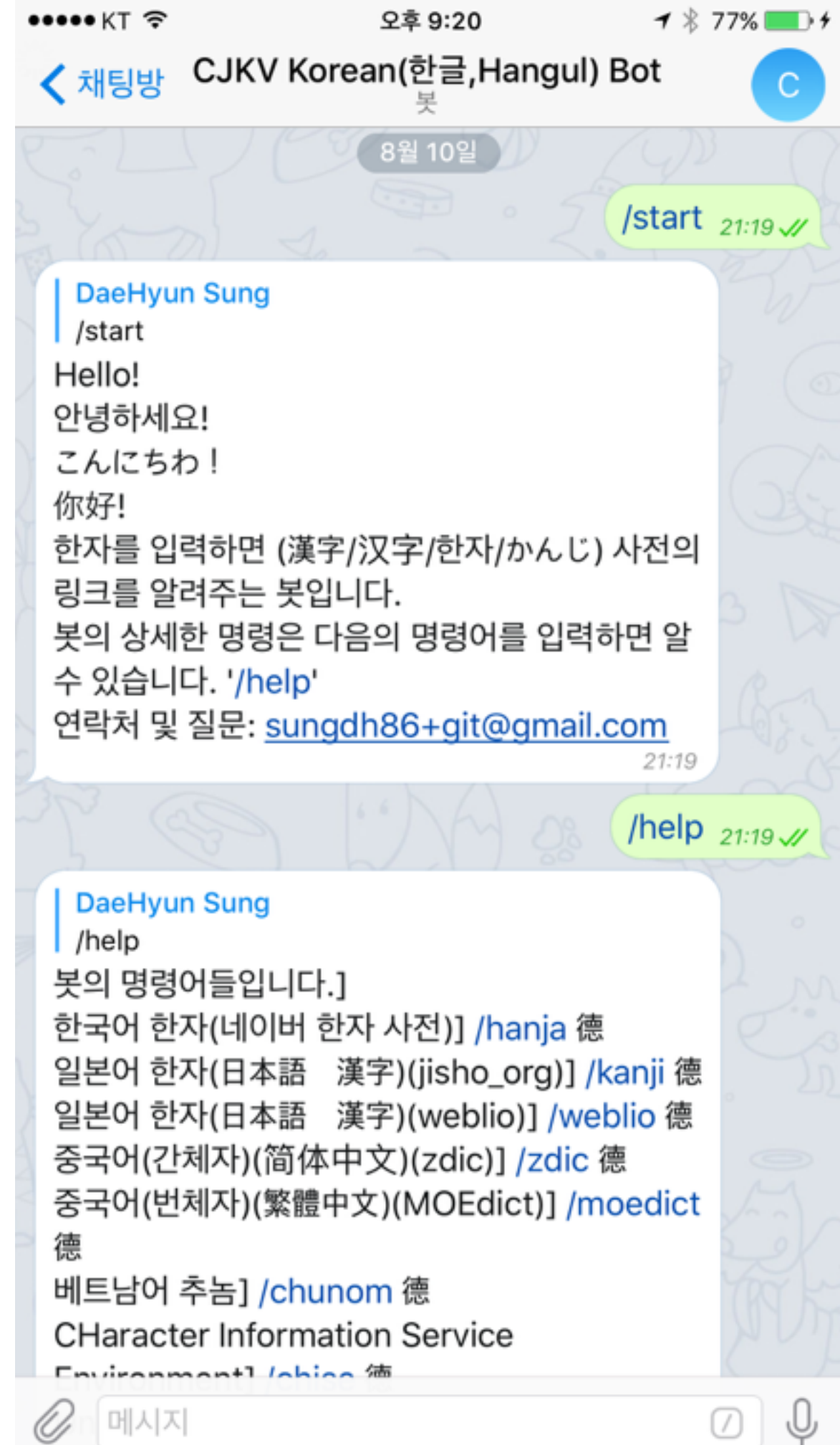
<https://twitter.com/unicode/status/722829746233409536>

德 성대현(成大鉉, DaeHyun Sung) ( @studioego )

<http://www.unicode.org/consortium/adopted-characters.html#b5FB7>

# Telegram Bot

- Telegram Bot (English)
  - @CJKVBot  
<https://telegram.me/cjkvBot>
- 텔레그램 봇 (한국어)
  - @CJKV\_korean\_bot  
[https://telegram.me/CJKV\\_korean\\_bot](https://telegram.me/CJKV_korean_bot)



# Chinese Character


- Chinese Character
- 漢字/汉字
- ㄏㄢˋ ㄓˋ /hànzì
- 한자
- かんじ
- hán tự

ㄏㄢˋ ㄗㄧˋ ㄨㄣˊ ㄏㄨㄚˋ ㄑㄩㄢ  
 漢字文化圈

hàn zì wén huà quān  
 汉字文化圈

hon<sup>3</sup> zi<sup>6</sup> man<sup>4</sup> faa<sup>3</sup> hyun<sup>1</sup>  
 漢字文化圈

한 자 문 화 권  
 漢字文化圈  
han ja mun hwa gwon

vòng vǎn hóa chũ hán  
 綏文化孺漢

かん じ ぶん か けん  
 漢字文化圈  
kan ji bun ka ken

Reference:


[https://commons.wikimedia.org/wiki/File:漢字文化圈%EF%BC%8F汉字文化圈\\_·\\_한자\\_문화권\\_·\\_Vòng\\_văn\\_hóa\\_chữ\\_Hán\\_·\\_漢字文化圈.jpg](https://commons.wikimedia.org/wiki/File:漢字文化圈%E2%BC%8F汉字文化圈_·_한자_문화권_·_Vòng_văn_hóa_chữ_Hán_·_漢字文化圈.jpg)

# Chinese Characters in Korea

- 한국어 (Korean)
  - 안녕하세요
  - 安寧하세요
- English: Hello
- 日本語 (Japanese) : こんにちは
- 漢語/汉语 (Chinese): 你好

same

# Chinese Characters in Korea

- 한국인  
한국의 국적을 가진 사람
  - 韓國人  
韓國의 國籍을 가진 사람
  - Korean  
People who have a Korean Nationality.
- 
- same



# Chinese Characters in Korea

대한민국 헌법  
제1장 총강  
제1조

① 대한민국은 민주공화국이다.

② 대한민국의 주권은 국민에게 있고, 모든 권력은 국민으로부터 나온다.

大韓民國 憲法  
第1章 總綱  
第1條

① 大韓民國은 民主共和國이다.

② 大韓民國의 主權은 國民에게 있고, 모든 權力은 國民으로부터 나온다.

same

Constitution of the Republic of Korea  
CHAPTER I . General Provisions  
Article 1 [Democracy]

(1) The Republic of Korea shall be a democratic republic.

(2) The sovereignty of the Republic of Korea shall reside in the people, and all state authority shall emanate from the people.

## 서울 1000萬 인구 25년만에 깨졌다

2030은 줄고 60대 이상 늘어  
지역내 성장률도 계속 둔화

서울의 노령화가 급속히 진행되면서 현재 1100만명 고지를 넘었던 서울시(市) 인구가 1000만명 아래로 떨어졌다. 이는 지난 1988년 말 기준으로 서울시 등록 인구가 1000만명을 돌파한 이후 25년 만이다.

10일 서울시에 따르면 작년 12월 말 기준 서울의 주민등록 인구는 1014만 3645명으로 나타났다. 하지만 실제로 서울에 살고 있는 지역구가 확인되지 않은 '거주(居住) 불명 등록자' 15만3003명을 제외한 총거주자 수는 998만9672명으로 집계됐다.

서울에 사는 내국인 수는 지난 1992년에 1000만9000명으로 최고(最上)를 기록했다. 이후 감소 추세가 이어지고 있다. 인구가 줄어드는 주된 이유로는 집값 등 주거에 드는 비용이 크게 오르면서 기존 주민들이 경기도와 인천·충남 등으로 대거 이탈했기 때문으로 분석된다.

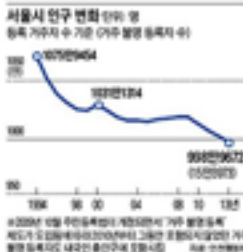


'수상한 그녀'

점점 세진 흥행 돌풍 ▶ A8면

발행연수

A36-B12



최근 통계청이 발표한 '2013년 국내 인구 이동 통계'에 따르면 지난해 전국적으로 가장 많은 사람이 빠져나간 지역이 서울이다. 서울시는 전입자에서 전출자 수를 뺀 순(純)이출 인구가 10만1000명에 달했다. 특히 전입신고율 비탈음으로 서울에서 빠져나간 이들이 어디로 갔는 지 살펴보니 전체의 60%가 경기도로 옮겨간 것으로 조사됐다.

인구 자체가 줄어드는 것도 문제지만 노년층 비율이 늘어나는 노령화(老齡化)가 빠르게 진행되는 것이 더 큰 문제라는 지적도 있다. 현재 60대 이상이 전체 서울 인구에서 차지하는 비율은 지난 2010년 14.6%에서 지난해 16.6%를 기록했다. 3년 새 무려 2%포인트가 증가했다.

서울에서 만 65세 이상 노인 인구는 지난 2010년 말 이미 100만명을 넘어섰고 오는 2027년엔 이보다 2배 많은 200만명을 돌파할 것으로 예상된다.

임지혜·이재훈 기자 A19면에 계속

## 北, 5만~10만달러 이상 資産家 24만명



2부: (2)호 개혁·개방 특주소

북한의 대외무역과 시장경제 규모가 커지면서 5만~10만달러 이상 재산을 보유한 자산가(資産家)가 24만명가량 되는 것으로 추정됐다. 북한 시장에 밝은 정파 소식통은 10일 "북한의 각종 경제 정세를 종합해 분석한 결과 과거 출신 성분을 중시

작년 경제개발구 14곳 만들면서  
달러·위안화 등 外貨유동 활발화

했던 태서 벗어나 개인이 사(私)경제 능력에 따른 새로운 계층 분화가 이뤄지고 있다"며 이렇게 말했다. 회고 자산가 중에는 중앙당 건학나 외관, 외화발에 상사원, 시장에서 돈을 번 '돈주(金主·金主)' 등이 많다고 이 관계자는 전했다.

한편 북한은 작년 사·도별로 '경제개발구' 14개를 만들면서 이 지역에서 미국 달러화와 중국 위안화, 유로화 등 외화(外

貨)의 유통 및 결제를 합법화하는 조치를 내린 것으로 확인됐다. 전국적 범위에서 외화 유통을 합법화한 것은 이번이 처음으로 북한의 시장화를 촉진할 수 있을지 주목된다.

정부 고위 관계자는 "북한이 작년 대외 개방 기본법으로 제정한 '경제개발구법' 제4조는 '경제개발구에서 유통 화폐와 결제 화폐는 조선 원 또는 장래전 화폐로 한다'고 규정했고, 제4조는 '경제개발구에서는 외화를 자유롭게 반출·입할 수 있다'고 했다"고 밝혔다. 북한법 전문가인 유육 변호사(법무법인 태평양)는 "제4

조에서 '장래전 화폐'는 곧 외화를 의미한다"며 "전국 단위에서 합법적인 외화 유통이 가능해졌다는 의미"라고 말했다. 북한은 대중식 시장에서 외화를 결제 수단으로 광범위하게 사용하고 있지만 합법적 유통은 개성공단, 나선각구 등 '철조망'으로 제한된 지역에서만 허용됐다.

김병현 서울대 교수는 "외화 유통이 자유화되면 투자 환경은 개선되겠지만 당국의 잦은 정책 변경 등 국가적 리스크가 여전히 실제 현금을 투자할 기업은 많지 않을 것"이라고 했다.

황대진 기자 기사 A4-5면



모태범 아들은 4위- 네덜란드, 男500m 금은동 박봉이 모태범(25·대전광역시)이 10일 아를레르 아레나에서 열린 2014 소치 동계올림픽 남자 스피드스케이팅 500m 1차 시기에서 역주를 펼치고 있다. 이 종목 동메달 2연패에 도전했던 모태범은 1-2차 시기 합계 1분09초88로 4위를 차지했다. 이철 말더러(1분08초31), 안 스메이팅(1분08초30), 로날트 빌더(1분08초40)도 네덜란드 선수들이 금·은·동메달을 휩쓸었다. 기사 A2면

▶ 기사 A2-3-24-26-27면

## "북한의 중산층 최근 크게 늘면서 갈수록 소비지향적"

아프트 북한출퇴기업인 대표

북한에 진출한 유럽 연합(EU) 기업들의 모임인 '평양 유럽 기업가 연합회'의 초대 대표를 지낸 알렉스 아프트(Alex·사진)씨는 지난달 28일 본지 아레일 인터뷰에서 "최근 북한 사회의 가장 큰 변화는 각종 비즈니스에서 '놀라울 정도로 열광' 모습을 보이는 중산층이 크게 성장하고 이들이 갈수록 소비 지향적 경향을 띠는 것"이라고 말했다.

스위스 출신인 아프트씨는 북한에서 의약품 제조 사업을 해왔다. 또 일종의 MBA 과정인 '평양 비즈니스 스쿨'을 설립해 시장경제를 가르쳤고, 2012년 '북한에 간 자본주의자-은둔의 왕국에서의 7년'이란 책도 펴냈다.

아프트씨는 "북한 주민들은 1990년대 '고난의 행군' 때 공공 배급 제도가 붕괴되면서 시장경제에 의존할 수밖에 없었다"며 "북한 주민들은 그해 이미 자본주의의 '견습생 생활(apprenticeship)'을 경험했다"고 했다.

황대진 기자, 이기훈 인천기자 A5면에 계속



## sochi 2014

## 왼쪽 무릎에 물이 차지만... 오늘밤 光速 레이스 시작

이상화, 올림픽 500m 2연패 도전

"박고에 갔다 오니 무릎에 물종이 오기 시작했다. 나는 약을 바르고 기면이 뇌무였다. 이제부터는 무릎이 다치지 않도록 조심해야겠다." "빙속 여제" 이상화(25·서울시청)가 은석초등학교 6학년 때 일가게 적은 내용이다. 다리를 굽혔다 폈다 하는 동작을 반복하는 스피드스케이팅 선수에게 무릎 부상은 숙명이다. 11일 오후 9시 45분(한국 시각)부터

자 스피드스케이팅 500m 올림픽 2연패(銀牌)에 도전하는 이상화의 왼쪽 무릎엔 현재 물이 차 있다. 이상화의 아버지 이우근(57)씨는 "(이)상화가 '벤쿠버올림픽'이 끝나고 무릎 수술을 받으려고 했지만 시즌이 계속 진행돼 일단 미룬 상태"라며 "정기적으로 치료를 받고 있다"고 말했다. 이 상화의 치료를 맡았던 조상연 하남병원 원장은 "벤쿠버올림픽 전부터 무릎에 물이 약간 차 있는 상태라 꾸준히 치료를 해 왔다"며 "소치올림픽이 끝나면 무릎 수술을

결도해야 할 것"이라고 말했다.

무릎에 물이 차는 증상은 연골이나 연골판이 손상됐을 때 약해진 부위의 세포들이 자가방어(自衛)를 위해 물을 배출해 생기는 현상이다. 축구 스타 박지성(PSV 에인트호벤)이 2011년 서른 살의 나이에 왼쪽 대표팀 은퇴를 선언한 결정적 이유도 무릎에 물이 차 부어오르는 증상 때문이었다. 동방석·오유경 기자 A2면에 계속

▶ 1차 레이스 21시 45분 ▶ 2차 23시 34분



"빙속 여제" 이상화가 은석초등학교 6학년 때 쓴 일기장. '무릎'이란 제목의 2001년 8월 15일자 일기에 '의자에 앉는 순간 무릎을 구부리니까 무릎이 아프고 소리가 나고, 천천히 누르면 아프고 이상한 느낌이 든다'고 적었다.

八面輝

하 대령원, "공공 노조 개혁 제안을 남 않을 것" 대령원보다 세다고 생각하는 노조들이 어떤 표정 지을까?

○ 서울 인구, 20~30대 줄어들면서 1000만 밑으로 떨어지. 대한민국 수도, 늑지 말고 성숙해져.

○ 미셀 오마레, "미국의 열대" 역할 어떤 머물며 남한 인기 주를. '장풍(張風)' 여사'의 이미지 관리?

조선일보, 朝鮮日報, The Chosun Ilbo  
2014-02-11

One of the major newspapers in South Korea.

# CJKV Dictionary Link

- **한국어 한자사전(漢字辭典, Korean-Chinese Character Dictionary)**
  - 네이버 한자사전(Naver Hanja Dictionary)  
<http://hanja.naver.com/>
- **Japanese Dictionary**
  - Weblio <http://ejje.weblio.jp>
  - Jisho <http://jisho.org>
- **Vietnamese Chữ Nôm Dictionary**
  - Chu nom <http://www.chunom.org>

# CJKV Dictionary Link

- Chinese Dictionary
  - ZDict (简体中文) <http://www.zdic.net>
  - MOE Dict (**M**inistry **O**f **E**ducation, Taiwan's Chinese Dictionary) (繁體中文)  
<https://www.moedict.tw>



# CJKV Dictionary Link

- **Glyph Wiki**

- English <http://en.glyphwiki.org>
- 한국어(Korean) <http://ko.glyphwiki.org>
- 日本語(Japanese): <http://glyphwiki.org>
- 简体中文(Simplified Chinese)  
<http://zhs.glyphwiki.org>
- 繁體中文(Traditional Chinese)  
<http://zht.glyphwiki.org>

# CJKV Dictionary Link

- **Unicode Consortium**
  - Unihan database lookup
  - <http://www.unicode.org/charts/unihan.html>
- **Chise (CHaracter Information Service Environment)**
  - <http://www.chise.org/ids-find>

# Unicode

- Unihan Database
  - <http://www.unicode.org/charts/unihan.html>
- Code Point
  - 德 - U+5FB7
    - dak1, 덕, dé, OSHIE, TOKU, đức
  - 成 - U+6210
    - cing4 seng4 sing4, 성, chéng, NARU TAIRAGERU NASU, SEI JOU, thành

U+5FB7

Chinese Character:德|CodePoint: U+5FB7|field: kCantonese|content: dak1

kCantonese = CodePoint: U+5FB7|field: kCantonese|content: dak1

U+5FB7

Chinese Character:德|CodePoint: U+5FB7|field: kDefinition|content: ethics, morality, virtue

kDefinition = CodePoint: U+5FB7|field: kDefinition|content: ethics, morality, virtue

U+5FB7

Chinese Character:德|CodePoint: U+5FB7|field: kHangul|content: 덕

kHangul = CodePoint: U+5FB7|field: kHangul|content: 덕

U+5FB7

Chinese Character:德|CodePoint: U+5FB7|field: kHanyuPinlu|content: dé(97)

kHanyuPinlu = CodePoint: U+5FB7|field: kHanyuPinlu|content: dé(97)

U+5FB7

Chinese Character:德|CodePoint: U+5FB7|field: kHanyuPinyin|content: 20841.170:dé

kHanyuPinyin = CodePoint: U+5FB7|field: kHanyuPinyin|content: 20841.170:dé

U+5FB7

Chinese Character:德|CodePoint: U+5FB7|field: kJapaneseKun|content: OSHIE

kJapaneseKun = CodePoint: U+5FB7|field: kJapaneseKun|content: OSHIE

U+5FB7

Chinese Character:德|CodePoint: U+5FB7|field: kJapaneseOn|content: TOKU

kJapaneseOn = CodePoint: U+5FB7|field: kJapaneseOn|content: TOKU

U+5FB7

Chinese Character:德|CodePoint: U+5FB7|field: kKorean|content: TEK

kKorean = CodePoint: U+5FB7|field: kKorean|content: TEK

U+5FB7

Chinese Character:德|CodePoint: U+5FB7|field: kMandarin|content: dé

kMandarin = CodePoint: U+5FB7|field: kMandarin|content: dé

U+5FB7

Chinese Character:德|CodePoint: U+5FB7|field: kTang|content: dək

kTang = CodePoint: U+5FB7|field: kTang|content: dək

U+5FB7

Chinese Character:德|CodePoint: U+5FB7|field: kVietnamese|content: đức

kVietnamese = CodePoint: U+5FB7|field: kVietnamese|content: đức

U+5FB7

Chinese Character:德|CodePoint: U+5FB7|field: kXHC1983|content: 0224.010:dé

kXHC1983 = CodePoint: U+5FB7|field: kXHC1983|content: 0224.010:dé



U+6210

Chinese Character:成|CodePoint: U+6210|field: kCantonese|content: cing4 seng4 sing4

kCantonese = CodePoint: U+6210|field: kCantonese|content: cing4 seng4 sing4

U+6210

Chinese Character:成|CodePoint: U+6210|field: kDefinition|content: completed, finished, fixed

kDefinition = CodePoint: U+6210|field: kDefinition|content: completed, finished, fixed

U+6210

Chinese Character:成|CodePoint: U+6210|field: kHangul|content: 성

kHangul = CodePoint: U+6210|field: kHangul|content: 성

U+6210

Chinese Character:成|CodePoint: U+6210|field: kHanyuPinlu|content: chéng(4462) cheng(17)

kHanyuPinlu = CodePoint: U+6210|field: kHanyuPinlu|content: chéng(4462) cheng(17)

U+6210

Chinese Character:成|CodePoint: U+6210|field: kHanyuPinyin|content: 21399.010:chéng

kHanyuPinyin = CodePoint: U+6210|field: kHanyuPinyin|content: 21399.010:chéng

U+6210

Chinese Character:成|CodePoint: U+6210|field: kJapaneseKun|content: NARU TAIRAGERU NASU

kJapaneseKun = CodePoint: U+6210|field: kJapaneseKun|content: NARU TAIRAGERU NASU

U+6210

Chinese Character:成|CodePoint: U+6210|field: kJapaneseOn|content: SEI JOU

kJapaneseOn = CodePoint: U+6210|field: kJapaneseOn|content: SEI JOU

U+6210

Chinese Character:成|CodePoint: U+6210|field: kKorean|content: SENG

kKorean = CodePoint: U+6210|field: kKorean|content: SENG

U+6210

Chinese Character:成|CodePoint: U+6210|field: kMandarin|content: chéng

kMandarin = CodePoint: U+6210|field: kMandarin|content: chéng

U+6210

Chinese Character:成|CodePoint: U+6210|field: kTang|content: \*zhieng

kTang = CodePoint: U+6210|field: kTang|content: \*zhieng

U+6210

Chinese Character:成|CodePoint: U+6210|field: kVietnamese|content: thành

kVietnamese = CodePoint: U+6210|field: kVietnamese|content: thành

U+6210

Chinese Character:成|CodePoint: U+6210|field: kXHC1983|content: 0135.080:chéng

kXHC1983 = CodePoint: U+6210|field: kXHC1983|content: 0135.080:chéng

# Dictionary URL Patterns

- Patterns
  - Example) 德
    - Chinese Character
      - <https://www.moedict.tw/德>
      - <http://hanja.naver.com/hanja?q=德>
    - Code Point
      - <http://www.zdic.net/z/jd/?u=5fb7>
      - <http://www.unicode.org/cgi-bin/GetUniHanData.pl?codepoint=5FB7>

# Development


- Python3
- Heroku (Free Dyno)
- Telegram python API

- <https://github.com/eternnoir/pyTelegramBotAPI>



# Create Telegram Bot



BotFather   
bot



DaeHyun  
/newbot

✓ 25.06.16



BotFather

25.06.16

Alright, a new bot. How are we going to call it? Please choose a name for your bot.



DaeHyun

✓ 25.06.16

CJKV Korean(한글, Hangeul) Bot



BotFather

25.06.16

Good. Now let's choose a username for your bot. It must end in `bot`. Like this, for example: TetrisBot or tetris\_bot.



DaeHyun

✓ 25.06.16

CJKV\_korean\_bot



BotFather

25.06.16

Done! Congratulations on your new bot. You will find it at [telegram.me/CJKV\\_korean\\_bot](https://telegram.me/CJKV_korean_bot). You can now add a description, about section and profile picture for your bot, see [/help](#) for a list of commands. By the way, when you've finished creating your cool bot, ping our Bot Support if you want a better username for it. Just make sure the bot is fully operational before you do this.

Get a Bot Token: xxxx2134:1234:xxx (example)

# Result



DaeHyun  
/moedict 德



CJKVBot  
DaeHyun Sung  
/moedict 德

德

CJK UNIFIED IDEOGRAPH-5FB7

<https://www.moedict.tw/德>

[www.moedict.tw](http://www.moedict.tw)

德 - 萌典

人類共同遵循的規範。品行、作風。恩澤、恩惠。心意、信念。德意志聯邦共和國的簡稱。姓。如漢代西域有德若國，後人以國為姓，簡姓德。感激、感恩。好的、善的。



DaeHyun  
/hanja 韓



CJKVBot  
DaeHyun Sung  
/hanja 韓

韓

CJK UNIFIED IDEOGRAPH-97D3

<http://hanja.naver.com/hanja?q=韓>

# Q&A, Reference



- Twitter @studioego
- sungdh86+git at gmail.com
- Github  
<https://github.com/studioego/CJKV/tree/master/unihan-dict-link>
- CJKV Information Processing, 2nd Edition  
Written by Ken Lunde
- PyCon APAC 2016 Seoul Lightning Talk(Korean)  
<http://www.slideshare.net/studioego/python-64911528>