

# <Emory-Yonsei Entity Guideline>

## A. General Principles

1. If an entity exists only in a particular language due to language differences, it is tagged only in that language without mapping.

### [Examples]

If there is a sentence in Korean that says, "*There was not even one ~*", "One" is tagged as "CARDINAL". While in English, 'any' is used instead of 'one', and so there is no cardinal expression in English.

2. If Korean and English taggings varies depending on context, English tagging is used.

### [Examples]

If a word "Korea" in Korean sentence contains meaning as 'Korean' and it is also expressed as "Korean" in English version of the sentence, the word 'Korea' in Korean sentence is tagged as "NORP", not 'GPE'.

3. When tagging classification is ambiguous as an entity derived from newly coined terms or the latest industry and etc, it basically followed English tagging.

### [Examples]

If 'AI(Artificial Intelligence)' is tagged as 'PRODUCT' in English, it is also tagged as 'PRODUCT' in Korean.

## B. Tag Set

### 1. Enamex

#### (1) **Person**: <tag = 'PERSON'>

Referring to a last name or full name of a particular person. It also includes nicknames for non-human creatures and characters in cartoons, dramas and movies.

\* Names of singers or celebrities and names of singer groups are included.

\* The positions or other additional expressions attached to the above persons shall be excluded. (e.g. Mister, Professor, etc)

#### [Examples]

(1) ENG: Soccer player Son Heung-min<PERSON> as a product model, is representative.

KOR: 축구선수 손흥민<PERSON>을 제품 모델로 기용한

(2) ENG: from Aristotle<PERSON> to Hegel<PERSON>, the philosophy of totality

KOR: 특히 아리스토텔레스<PERSON>에서 헤겔<PERSON>까지 이어진 본질주의

(3) ENG: It was Ryan<PERSON> who flew by plane because he wanted to see Deokmi<PERSON>.

KOR: 바로 덕미<PERSON>가 보고 싶어서 비행기를 타고 날아온 라이언<PERSON>

#### (2) **Norp**: <tag = 'NORP'>

It refers to nationality, religious groups and political groups (party).

#### [Examples]

(1) ENG: A few years ago, in an Italian<NORP> restaurant

KOR: 몇 년 전, 한 이탈리안<NORP> 레스토랑에서

(2) ENG: for Christian<NORP> culture to permeate through soccer

KOR: 기독교<NORP> 문화가 스며들도록 전도하는 특수사역이었다.

(3) ENG: the majority of Democratic Party of Korea<NORP>

KOR: 민주당<NORP> 다수의 의원들이 이 원내대표를 민 것 자체가

### (3) Facility: <tag = 'FAC'>

Objectives referring to facilities include buildings, airports, highways and bridge names.

#### [Examples]

- (1) ENG: Under the Margaret Bridge<FAC> near the accident site

KOR: 사고 지점 인근 마가렛 다리<FAC> 아래에는 희생자들을 애도하고

- (2) ENG: Not long ago, the Empire State Building<FAC>

KOR: 얼마 전엔 엠파이어 스테이트 빌딩<FAC>이

### (4) Organization: <tag = 'ORG'>

It refers a community / group of people gathered together. For example, the name of the company, the name of the school, and the name of the sports team.

#### [Examples]

- (1) ENG: The Kookmin Ilbo<ORG> reported the analysis that many companies

KOR: 국민일보<ORG>는 야간 유람선 사업에 많은 업체들이 우후죽순 뛰어들면서

- (2) ENG: but Barcelona<ORG>'s midfield is saturated

KOR: 바르셀로나<ORG>의 중원은 포화상태다

- (3) ENG: France 24<ORG> and the New York Times<ORG> reported that

KOR: 프랑스 24<ORG>, 뉴욕타임스<ORG> 등은 현지에서 전사한

### (5) Geolocation: <tag = 'GPE'>

An object referring to a place or location, including the name of a country and the name of an administrative district, such as a city or state.

### [Examples]

(1) ENG: the ongoing crisis in Japan<GPE> and the U.S.<GPE>, and challenges

KOR: 일본<GPE>과 미국<GPE>에서 현재 진행 중인 위기와 전 세계가 해결해야 할 과제

(2) ENG: surfing, relaxing and shopping at the same time while living in New York<GPE>.

KOR: 뉴욕<GPE>에 거주하면서 해변 물놀이와 서핑, 휴식과 쇼핑을

### (6) Location: <tag = 'LOC'>

Refers to the name of a place or location that does not belong to GPE. It also includes expressions covering the entire location of mountains, rivers, ocean names and Europe, Asia, etc.

### [Examples]

(1) ENG: In the Danube river<LOC> after the disaster, citizens are gathering in honor of

KOR: 참사 이후 다뉴브강<LOC>에서는 희생자들을 추모하는 시민들이 모이고 있다.

(2) ENG: It was hard for Western Europe<LOC> and

KOR: 신문사가 자체 검열하는 다른 나라의 모습을 서유럽<LOC>이나

(3) ENG: if the network is actually formed begins in Europe<LOC>.

KOR: 유럽<LOC> 등에서 네트워크 구축이 본격화되면 점유율 변화가 올 것이라는

### (7) Product: <tag = 'PRODUCT'>

Vehicles, Weapons, foods. IT services (including SNS) and medicine names are included.

### [Examples]

(1) ENG: iPhones<PRODUCT> have decreased greatly as the

KOR: 아이폰<PRODUCT> 판매가 크게 줄어든 것처럼

(2) ENG: Google Glass<PRODUCT> for general consumers in 2013.

KOR: 구글 글라스<PRODUCT>를 출시하며 한때 혁신적인 제품으로 관심을 모았다.

(3) ENG: online music service Melon<PRODUCT> had been interrupted

KOR: 온라인 음악서비스 멜론<PRODUCT>이 중단됐었고

### (8) Event: <tag = 'EVENT'>

It means an official or widely known event, war, exhibition. Official events include ministerial meetings, general elections, presidential elections, exams (SAT), and prayers (U.S. national breakfast prayer). Social phenomena also include (Brexit) for widely known events.

#### [Examples]

(1) ENG: Cannes International Film Festival<EVENT> for the film

KOR: 칸 국제영화제<EVENT>에서 커리어의 정점에 오른

(2) ENG: provides live PGA Tour<EVENT> coverage as well as video-on-demand.

KOR: PGA 투어<EVENT> 생중계는 물론 주문형비디오

(3) ENG: the International Economic Forum<EVENT> in

KOR: 열리는 국제경제포럼<EVENT>에

### (9) Work of Art: <tag = 'WORK\_OF\_ART'>

Titles of books, songs, TV programs and art pieces. Title of games, awards, theories, records are included.

#### [Examples]

(1) ENG: "Ugly Duckling<WORK\_OF\_ART>" was a book that I often went to.

KOR: '미운오리새끼<WORK\_OF\_ART>'도 자주 손이 가던 책이었다.

(2) ENG: for the film "Parasite<WORK\_OF\_ART>".

KOR: 영화 '기생충<WORK\_OF\_ART>'으로

(3) ENG: "Spring Night<WORK\_OF\_ART>" had high expectations before the show

KOR: ‘봄밤<WORK OF ART>’은 파격 편성, 멜로 거장

**(10) Law: <tag = ‘LAW’>**

Named documents made into laws.

**[Examples]**

(1) ENG: completely denied the 1965 Claims Agreement<LAW>, the cornerstone of

KOR: 관계의 주춧돌인 1965 년 청구권협정<LAW>을 전면 부인한 판결을

(2) ENG: in accordance with the Framework Act on Science and Technology<LAW>.

KOR: 현재 각 부처가 진행하는 연구개발 사업은 기본적으로 과학기술기본법<LAW>

**(11) Language: <tag = ‘LANGUAGE’>**

Any named language.

**[Examples]**

(1) ENG : the quality of classes centered on Korean<LANGUAGE>

KOR : 국어<LANGUAGE>나 수학 교과를 중심으로 수업의 질을 높이기 위해

(2) ENG: often translated into "assemblage" in English<LANGUAGE> and "deployment" in

Korean<LANGUAGE> and Japanese<LANGUAGE>

KOR: ‘agencement’은 흔히 영어<LANGUAGE>로 ‘assemblage’, 한국어<LANGUAGE>와

일본어<LANGUAGE>로는 ‘배치’로 번역되는데

## **2. Timex**

### **(1) Date: <tag = 'DATE'>**

Absolute or relative dates or periods. The period should last longer than 'TIME'. General expressions of dates are included too such as 'few months', 'that day', 'Next season' and 'First quarter'.

#### **[Examples]**

(1) ENG: Electronic retailers spent a busy May<DATE> as consumers rush to buy

KOR : 에어컨 구매를 서두르면서 전자 유통업체들은 분주한 5 월<DATE>을 보냈다.

(2) ENG: goes well even in the late summer<DATE> when the weather starts to get cool.

KOR: 얇은 외투처럼 걸칠 수 있어 더위가 가시기 시작하는 늦여름<DATE>에도 어울리는

(3) ENG: for diving until June 2<DATE> due to the high flow rate

KOR: 회의 끝에 유속이 빨라 6 월 2 일<DATE>까지

### **(2) Time: <tag = 'TIME'>**

'Time' tag indicates time expressions smaller than a day. This tag includes certain time indication, amount of time or any other expressions related to time. Even though an entity does not have numeral expressions but only words related to time (for instance, 'noon'), the words are tagged as 'Time'.

#### **[Examples]**

(1) ENG: around 12:45 pm<TIME>, had to go in and out of water every

KOR: 낮 12 시 45 분<TIME>께 투입된 잠수부

(2) ENG : about a five-minute<TIME> distance from Hanmaum Hall

KOR : 한마음회관에서 5 분<TIME> 정도 거리이고

### **3. Numex**

#### **(1) Percent:** <tag = 'PERCENT'>

Percentage expressions with % symbol or the word 'percent'.

##### **[Examples]**

(1) ENG: could impose high tariffs of 25 percent<PERCENT> on

KOR: 자동차에 25<PERCENT>에 고율 관세를 매길 수 있다고 경고한 바 있다.

(2) ENG: The export order index fell 2.7 percent<PERCENT> from a month

KOR: 2.7<PERCENT> 하락했고 수입주문지수 역시

#### **(2) Money:** <tag = 'MONEY'>

Monetary values including units. \* Bitcoin is not included

##### **[Examples]**

(1) ENG: exchange rate fell to the 108 yen<MONEY> level.

KOR: 엔화 가치가 오르면서 엔·달러 환율은 108 엔<MONEY>대로 떨어졌다.

(2) ENG: they had made an unfair claim of 87.73 million won<MONEY> for

KOR: 8773 만원<MONEY>을 부당청구한 사실을 밝혀낸 뒤

#### **(3) Quantity:** <tag = 'QUANTITY'>

Measurements as of weights or distances such as km, kg and etc.

##### **[Examples]**

(1) ENG: The thickness of the church conference record was about 3cm<QUANTITY>.

KOR: 한 회 분량의 당회록 두께가 3 cm<QUANTITY> 정도였다.

(2) ENG: about 170,000 tons<QUANTITY>, of which 140,000 tons<QUANTITY> are from



KOR: 생산량은 약 17 만 톤<QUANTITY>이며, 이 중 14 만 톤<QUANTITY>이

#### (4) **Ordinal**: <tag = 'ORDINAL'>

All ordinal numbers such as first and second.

##### [Examples]

(1) ENG: it is difficult to score first<ORDINAL> place in the test

KOR: 놀 거 다 놀고 시험에서 1 등<ORDINAL> 하기는 어려운 일이다.

(2) ENG: that ended with the 16th<ORDINAL> episode on

KOR: 16 화<ORDINAL> 를 끝으로 종영했다.

#### (5) **Cardinal**: <tag = 'CARDINAL'>

##### [Description]

Numerical terms not categorized in other categorizations. Numbers that indicate ages are included.

##### [Examples]

(1) ENG: to voluntarily correct four<CARDINAL> clauses:

KOR: 사업자의 포괄적 면책, 부당한 재판관할 합의 등 4 개<CARDINAL> 조항은

(2) ENG: As I said, there are hundreds<CARDINAL> of rescue and search

KOR: 방금 말씀드린 대로 수백명<CARDINAL>의 구조·수색 인력이 현장에

(3) ENG: killing 7<CARDINAL> people and leaving 19<CARDINAL> others missing.

KOR: 침몰해 7 명<CARDINAL>이 숨지고 19 명<CARDINAL>이 실종됐다.