ACE (Automatic Content Extraction)
Arabic Annotation Guidelines for Entities

Version 7.4.2 June 13th, 2008

Linguistic Data Consortium

http://www.ldc.upenn.edu/Projects/ACE/
6 Nickname Metonymy ........................................................................................................42
   6.1 Capital City or Government Seat (FAC) Names standing in for Country’s Government ........................................................................................................43
   6.2 City name for Sports Team ..................................................................................43
7 Cross-Type Metonymy .................................................................................................43
Appendix: ......................................................................................................................45
   FAC.Building .............................................................................................................45
   FAC.Subarea .............................................................................................................45
   GPE.Cluster .............................................................................................................45
   GPE.Continent .......................................................................................................45
   GPE.County .............................................................................................................46
   GPE.Nation .............................................................................................................46
   GPE.Pop ..................................................................................................................46
   Gaza .......................................................................................................................46
   GPE.Special ............................................................................................................47
   GPE.State ................................................................................................................47
   LOC.Land-Reg .......................................................................................................47
   LOC.Reg-Gen .........................................................................................................47
   LOC.Reg-Int ............................................................................................................48
   LOC.Water ..............................................................................................................48
   Land Areas Named by Water References ................................................................48
   ORG.Non ................................................................................................................49
   Palestine ...............................................................................................................49
   NOT TAGGABLE ....................................................................................................49
1 Introduction
The Entity Detection task requires that selected types of entities mentioned in the source data be detected, their sense disambiguated, and that selected attributes of these entities be extracted and merged into a unified representation for each entity.

Basic Concepts
An entity is an object or set of objects in the world. A mention is a reference to an entity. Entities may be referenced in a text by their name, indicated by a common noun or noun phrase, or represented by a pronoun. For example, the following are several mentions of a single entity:

Name Mention:

Nominal Mention:

Pronoun Mentions:

Entities are limited to the following five types:

- Person - Person entities are limited to humans. A person may be a single individual or a group.
- Organization - Organization entities are limited to corporations, agencies, and other groups of people defined by an established organizational structure.
- GPE (Geo-political Entity) - GPE entities are geographical regions defined by political and/or social groups. A GPE entity subsumes and does not distinguish between a nation, its region, its government, or its people.
- Location - Location entities are limited to geographical entities such as geographical areas and landmasses, bodies of water, and geological formations.
- Facility - Facility entities are limited to buildings and other permanent man-made structures and real estate improvements.

For each entity, the annotation records the type of the entity (PER, ORG, GPE, LOC, FAC), subtype, class, and all the textual mentions of that entity.

2 Text to Annotate
Only material between <TEXT> and </TEXT> tags is to be annotated.
Metadata imported from other sources is not part of the source signal from which ACE systems are presumed to extract information. Therefore no information is to be extracted from such metadata. Hence, in newswire documents, material in headlines and slug sections is not to be tagged. In broadcast data and telephone speech data, only the transcribed speech is to be tagged. Added information, i.e. <TURN> tags or speaker identification tags, is NOT to be tagged. In weblog and newsgroup files, we do not tag the poster name.

In the transcribed speech data, we don't tag any hesitation except if the word is transcribed as a complete word. The following example explains this idea:

<table>
<thead>
<tr>
<th>The ca- a- a- r</th>
<th>Not taggable</th>
</tr>
</thead>
<tbody>
<tr>
<td>The ca- a- a- the car</td>
<td>Taggable</td>
</tr>
</tbody>
</table>

3 Entity Types and Subtypes
Bracket indicates the extent of a mention while bold with underline indicates the head of a mention.

3.1 Persons (PER)
Each distinct person or set of people mentioned in a document refers to an entity of type Person. For example, people may be specified by name ("سعد الحريري"), occupation ("the teacher"), family relation ("والد"), pronoun ("هو"), etc., or by some combination of these. Dead people and human remains are to be recorded as entities of type Person. So are fictional human characters appearing in movies, TV, books, plays, etc.

[اللقب السياسي لهذا العام]

[The political cat of the year]

[The Turkish forces attacked Iraq]

[Some students]

[Three Israeli soldiers]
There are a number of words that are ambiguous as to their referent. For example, nouns, which normally refer to animals or non-humans, can be used to describe people. If it is clear to the annotator that the noun refers to a person in a given context, it should be marked as a Person entity. This is common in weblog files.

3.1.1. Subtypes for Person
We will further classify Person entities with the following subtypes.

**PER.Individual**
If the Person entity refers to a single person, tag it as PER.Individual.

**PER.Group**
If the Person entity refers to more than one person, tag it as PER.Group unless the group meets the requirements of an Organization or a GPE described below. This will include family names and ethnic and religious groups that do not have a formal organization unifying them. Ethnic groups of people and religious groups that do not have a formal organization unifying them will be considered entities of Person.Group.

**PER.Indeterminate**
If from the context you can’t judge whether the Person entity refers to one or more than one person, tag it as PER.Indeterminate. We are not expecting to see such subtype in Arabic.
3.1.2 Titles, Honorifics, and Positions

In English, as well as in Arabic, titles and most honorifics precede the name. We will not consider these to be part of the name of a Person. We will annotate these as mentions in their own right. The parts of titles are taggable if they refer to entities. For example, in the string “وزير الخارجية المصري أبوالفیط”， there would be four mentions of three distinct entities. The two person mentions are co-referential, but the title Secretary of State is attributive while the proper name is tagged as Specific.

Saints and other religious figures

Religious titles such as saint, prophet, imam, or archangel are to be treated as titles.

References to “God” will be taken to be the name of this entity for tagging purposes. If it is used as a descriptor rather than a name, it will be considered a nominal mention. Note that capitalization information may not be available in speech transcripts.
3.1.3 Fictional characters, names of animals, and names of fictional animals
Names of fictional characters are to be tagged; however, character names used as TV show titles will not be tagged when they refer to the show rather than the character name.

[Batman] has become a popular icon

A lot of kids love [Bakar]

[Bokloz] teaches the kids how to behave

Names of animals are not to be tagged, as they do not refer to person entities. The same is true for fictional animals and non-human characters. These two examples do not yield mentions.

3.2 Organizations
Each organization or set of organizations mentioned in a document gives rise to an entity of type Organization. An Organization entity must have some formally established association. Typical examples are businesses, government units, sports teams, and formally organized music groups. Industrial sectors and industries are also treated as Organization entities.

3.2.1 Subtypes for Organizations
We will further classify Organization entities with the following subtypes. Organizations which do not fit into the subtypes defined below will not be tagged.

Government (GOV)
Government organizations are those that are of, relating to, or dealing with the structure or affairs of government, politics, or the state. The entire government of a GPE is excluded from this subtype and should be tagged GPE.ORG. Military organizations that are connected to the government of a GPE will be tagged as Government.
The Supreme Court in Washington] agreed to discuss the appeal

[Saudi police] arrested two terrorists

The United States closed down [its embassy in Doha]

[The Egyptian police] are looking for the suspects

Commercial (COM)
A commercial organization is an entire organization or a taggable portion of an organization that is focused primarily upon providing ideas, products, or services for profit.

Educational (EDU)
An educational organization is an entire institution or taggable portion of an institution that is focused primarily upon the furthering or promulgation of learning/education.

[Lebanese universities] include European and American academic programs

[Egyptian schools] are overcrowded

Entertainment (ENT)
Entertainment organizations are those whose primary activity is entertainment. This includes organizations such as Barnum and Bailey’s Circus and HBO, but
excludes provider giants such as Comcast and media conglomerates such as Disney and Time-Warner. These companies are all best annotated as commercial organizations.

[Kuwait’s **Miami group**] signed an agreement with Rotana productions

[**Naguib Al Rihany theater group**] presented the greatest comedy shows

**Non-Governmental Organizations (NonGov)**

Non-governmental organizations are those organizations that are not a part of a government or commercial organization and whose main role is advocacy, charity or politics (in a broad sense). By this definition, all of the following would be annotated as ORG.NonGov:

This subtype will include such diverse organizations as:

1. (Para-)Military Organizations:

   [**Al Aqsa Martyr’s Brigade**]

   [**Tamil Tigers**]

2. Political Parties:

   [**The Republican Party**]

   [**The Labor Party**]

3. Political Advocacy Groups and Think Tanks:

   [**ACLU**]
4. Professional Regulatory and Advocacy Groups:

[Arabic-Entities-Guidelines.doc V7.4.2

5. Charitable Organizations:

[The Red Cross]

[UNICEF]

[Doctors Without Borders]

6. International Regulatory and Political Bodies:

[UN]

[NATO]

[The World Bank]

The fighting between [the Taliban] and the Afghan protestors is taking place only a few kilometers from the borders of this former Soviet Republic in Asia.
Media (MED)

Media organizations are those whose primary interest is the distribution of news or publications, regardless of whether the organization is privately or publicly owned. This will include media companies such as Time Magazine, but will exclude media conglomerates such as Time-Warner which should be annotated as a commercial organization.

[Al Hayat newspaper]
[The New York Times]
[BBC]
[Sayedaty magazine]

[Official Sudanese TV] reported that an attack on Al Sonna Al Mohamadia mosque in Garafa village close to Om Dorman north of Khartoum had occurred.

We will continue our news from [VOA in Washington]

Religious (REL)

Religious organizations are those that are primarily devoted to issues of religious worship.

[The German Bishop's Union]
[The Vatican]

[اتحاد الأساقفة الألمان]
Medical-Science (SCI)
Medical-Science organizations are those whose primary activity is the application of medical care or the pursuit of scientific research, regardless of whether that organization is publicly or privately owned.

[Al Azhar administration]

Palestinian hospitals are suffering greatly as physicians tend to the seriously injured on a daily bases and are short on auxiliary hospital staff.

[Chinese Overseas Physics Society]

[The Nuclear Research Institute]

Sports (SPO)
Sports organizations are those that are primarily concerned with participating in or governing organized sporting events, whether professional, amateur, or scholastic. We will include groups whose sports are board games, card games, and games of chance in this category.

[Al Negma] beat [Shabab Al Sahel] yesterday in the Lebanese tournament

[The Philippine Olympic Committee]

[Saudi Soccer Federation]

3.2.2 Organization Subtype Trumping Rules
The collection of organization subtypes is hierarchical in nature. Some organizations will naturally fit into more than one category. The following
The diagram displays the hierarchy of organization subtypes. As a rule, we will assign the most specific type possible.

<table>
<thead>
<tr>
<th>Less Specific</th>
<th>More Specific</th>
</tr>
</thead>
<tbody>
<tr>
<td>REL</td>
<td></td>
</tr>
<tr>
<td>GOV</td>
<td>EDU SCI</td>
</tr>
<tr>
<td>NonGov</td>
<td>ENT SPO</td>
</tr>
<tr>
<td>COM</td>
<td>MED</td>
</tr>
</tbody>
</table>

**Exceptions to Trumping Rules**

**GPE military**
The military organizations connected to a GPE’s government will be tagged Government (GOV), for example: US Air Force.

**Media Conglomerates**
Big media conglomerates such as Disney and Time-Warner will be tagged Commercial (COM). The subsidiary media organizations owned by these companies will be tagged Media (MED).

**Medical Schools and Research Labs**
Medical schools will be tagged Educational. Specific labs and research institutions which primarily devote their attention to medical or scientific research will be tagged Medical/Science (SCI) even when they are attached to educational institutions.

**Soft Science Research Institutions**
Institutions whose primary activity is the study of social sciences will be tagged Non-Governmental (NonGov).
**Boy Scouts**
The Boy Scouts of America and similar organizations will be tagged Educational (EDU).

3.2.3 Organization Entities used in Person Contexts
Whenever an organization takes an action, there are people within or in charge of the organization that one presumes actually made the decision and then carried it out. Thus many organization mentions could be thought of as metonymically referring to people within the organization.

However, there seems to be little to be gained in the usual case by thus “reaching inside the organization” to posit a mention of a Person entity. It seems better to adopt the view that organizations can be agentive, and take action on their own. Only when something in the context draws particular attention to the people within the organization should a separate mention of a Person entity be marked.

Sets of people who are not formally organized into a unit are to be treated as a Person entity rather than an Organization entity. It is often difficult to tell the difference between Organization entities and collections of individuals tagged as PER.Group entities. Examples of organization-like nouns which are not organizations are “employees,” and “crew.” Although the members of a company or crew may work together in an organized and even hierarchical fashion, the groups are not organizations by themselves.

3.3 Geographical/Social/Political Entities (GPE)
Geo-Political Entities are composite entities comprised of a population, a government, a physical location, and a nation (or province, state, county, city, etc.). All mentions of these four aspects of a GPEs will be marked GPE and coreferenced.

![Diagram of Entity Types]

Arabic-Entities-Guidelines.doc V7.4.2
2008.06.13
In this sentence,

*The people of France welcomed the agreement.*

there are two mentions

- *[The people of France]* GPE
- *[France]* GPE

The mention of the population of France is marked GPE, rather than PER. These mentions would be coreference as they refer to different aspects of a single GPE.

Explicit references to the government of a country (state, city, etc.) are to be treated as references to the same entity evoked by the name of the country. Thus "the United States" and "the United States government" are mentions of the same entity. On the other hand, references to a portion of the government ("the Administration", "the Clinton Administration") are to be treated as a separate entity of type Organization, even if it may be used in some cases interchangeably with references to the entire government (compare "the Clinton Administration signed a treaty" and "the United States signed a treaty").

Sometimes the names of GPE entities may be used to refer to other things associated with a region besides the government, people, or aggregate contents of the region. The most common examples are sports teams:

*New York* defeated *Boston* 99-97 in overtime.

These are to be recorded as distinct entities, not as mentions of the GPE entity. Thus, in this example, both "New York" and "Boston" would evoke Organization entities. Please refer to 6.2 for more discussions.

### 3.3.1 Subtypes for GPEs

We will further classify GPE entities with the following subtypes. GPE entities which do not fit into the subtypes defined below will not be annotated.

**Continent**

Taggable mentions of the entireties of any of the seven continents: North America, South America, Antarctica, Europe, Asia, Africa, and Australia.

*The report included names of 24 countries mostly from* [Asia], Latin America, and Central America

*Europe* is trying to unite
Nation
Taggable mentions of the entireties of any nation.

There were two more accidents in [Kalyoubia Governorate north of Cairo] where a man and four women died.

State-or-Province
Taggable mentions of the entireties of any state, province, or canton of any nation.

Two Egyptians were injured by [Israeli] bullets

County-or-District
Taggable mentions of the entireties of any county, district, prefecture, or analogous body of any state/province/canton. Usually a County-or-District is bigger than a city, but smaller than a State-or-Province. Note in a lot of Arabic countries, there is not such infrastructure. You can find such mentions in files which reports news in other countries though.

This classification is according to US system, which may no apply to many other countries.

[Kingdom of Saudi Arabia] agreed to forgive part of Yemen's debt

[Kingdom of Saudi Arabia] agreed to forgive part of Yemen's debt

[Kingdom of Saudi Arabia] agreed to forgive part of Yemen's debt

[Kingdom of Saudi Arabia] agreed to forgive part of Yemen's debt
Population-Center
Taggable mentions of the entireties of any GPE below the level of County-or-District, including cities, villages in Arabic countries. Units that are smaller than villages are considered as LOC-Region-General. Please refer below for more information.

The ambassadors of Uzbekistan and Afghanistan met in the [Pakistani capital]

At least ten persons were injured in a road accident in [Ashmoun city]

The International Institute for Strategic Studies in [London] said that the international weapon trade has slowed down

The American Embassy spokesman said Tuesday that US Secretary of State Madeleine Albright will arrive in [Riyadh] later today

It seems that the Christmas celebrations in [Bethlehem] the town where Jesus was born, will be sad event this year

GPE-Cluster
Named groupings of GPEs that can function as political entities.

Eastern Europe
Western Europe
the European Union
the Middle East
Southeast Asia
Latin America
The West
The East
The Gulf should be tagged as a Location/Region /natural when it refers to the sea, but as GPE/ Clusters when referring to countries around the Gulf.

**Special**
A closed set of GPEs for which the conventional labels do not straightforwardly apply.

The [Palestinian] police arrested a Palestinian suspected of dealing with Israel.

The following entities are also tagged as GPE /Special

- *Northern Ireland*  
  ايرلندا الشمالية
- *Hong Kong*  
  هونغ كونغ
- *Taiwan*  
  تايوان

### 3.3.2 GPE-like Locations and Organizations

Incidental, non-political clusters of GPEs should be marked Location. Please refer to 3.4 for further discussion.

[جنوب الولايات المتحدة]  
[the southern United States]

Coalitions of governments, as well as the UN, are organizational bodies and should be marked Organization.

[الاتحاد الأوروبي] صوتوا لصالح تغليظ العقوبات على إيران

[حلف الأطلسي] يحذر من الإنتشار النووي

### 3.3.3 Formulaic GPE Constructions: Nested Region Names

A series of nested region names, such as "Provo, Utah" evokes one entity for each region. Thus "Provo, Utah" evokes one entity for the population center (with mention "Provo, Utah") and a second one for the state (with mention "Utah"). In Arabic, due to the absence of punctuation, there are several variants of this structure:
3.3.4 GPE Mention Roles

Annotators need to decide for each entity mention in the text which role (Person, Organization, Location, GPE) the context of that mention invokes. This judgment typically depends on the relations that the entity enters into.

In the examples above, the name “France” refers to a range of concepts. Annotators must select the Role which matches the function of the GPE mention.

The GPE role may be used in contexts that highlight the nation (or state or province or city, etc.) aspect of the GPE entity, as distinct from the government, populace, and location, but it may also be used in contexts referring to an indistinct amalgam of more than one of the aspects of a GPE (government, population, location, and nation).

- **GPE.ORG** - France signed a treaty with Germany last week.
- **GPE.PER** - France vacations in August.
- **GPE.LOC** - The world leaders met in France yesterday.
- **GPE.GPE** - France produces better wine than New Jersey.

The following sections give particular guidelines for frequently encountered cases, with examples.

**GPE.ORG**

GPE.ORG is used for GPE mentions that refer to the entire governing body of a GPE. It is important to differentiate between a part of the government (the executive branch, the courts) and the entire governing body. Below are some examples of contexts in which GPE.ORG should be used.

**Political Communication and Decision-making**
ORGs are responsible for decisions to take military actions. ORGs are also responsible for political communication events such as announcements, agreements, statements, denials, expressions of approval and disapproval, etc. So, if China agrees to something, China is a GPE.ORG.

Governments

While the entity type for governments is GPE, the role for governments should always be GPE.ORG.

But [the Russian government] and many politicians will be stridently critical of the United States if they believe they are being ignored.

Somali officials declared that [their government] concluded an agreement of cooperation with its former enemy Ethiopia

Note that when the government is related to person it is annotated as ORG. GOV

GPE.PER

As stated above, populations of a GPE are treated as GPE.PER. However, it is sometimes difficult to determine whether a reference to people is a reference to the population as a whole.

[The Japanese] have a considerable responsibility for the wars of the first half of the century

In this example, the phrase the Japanese may be interpreted as the population of Japan, or the government of Japan, or the Japanese military, or even some part of the Japanese population. If the annotator believes that the phrase in question refers to the population of the GPE, or most of the population of a GPE, then the annotation should be GPE.PER and the mention is a name mention. However, if the annotator believes the phrase refers to a group of people, then PER is the
assigned annotation and the mention is nominal because it does not refer to the name of a person. Example:

أفاد بنثر بأنه يمكن إنهاء الحظر الاقتصادي إذا سمح [العراقين] للمت装扮ين بأخذ واجبهم

Butler says those economic sanctions could be lifted soon if [the Iraqis] allow the inspectors to do their job.

**GPE.LOC**

GPE.LOC is used when a mention of a GPE entity primarily references the territory or geographic position of the GPE.

أفادت بأن عملية عزل الرئيس استراذا قد تستغرق

It was reported in [Riyadh City in Saudi Arabia] that 3 Britons were injured when their car was blown up

No one claimed responsibility for either of the two incidents that occurred one week after the bombing of the American Embassy building in [Sanaa]

U.S. warplanes flew over [Afghanistan]

In nested mentions of the form [child,[parent]], the parent GPE always takes a GPE role; the child’s role depends on context.

Dateline mentions of GPEs are given a location role.
GPE.GPE
GPE.GPE is used when more than one of the other GPE roles is being referenced at once or when no one role stands out in the context. Below are a few particular contexts in which GPE.GPE should always be used.

GPE Postmodifiers (Premodifiers in English)
Postmodifiers are inherently vague and difficult to decompose. For this reason, all GPE postmodifiers will be assigned the role GPE.GPE.

- [Israel] troops
- [New York] policemen
- [British] Prime Minister
- [U.S.] surveillance aircraft
- [Iraqi] flag

Clashes broke out on Wednesday in Jerusalem and its suburbs between Palestinians and the [Israeli] army.

Military Activity
Similarly, military activities like invasions, military strikes, bombings, etc. are considered to be acts carried out by and directed at entire nations (not distinguishable from the government, people and location of that nation) and therefore are associated with GPEs. Both the aggressors and the victims in these cases are marked GPE.GPE.
The city could have used some special protection in nineteen seventy-nine when the [Soviet Union] invaded Afghanistan.

**Activities Associated with GPEs**

Certain activities are associated with GPEs and therefore invoke a GPE role. For example, in a pro-Iraq rally, Iraq is assigned a GPE.GPE annotation. A rally is generally concerned with a nation as a whole, rather than exclusively a location or government.

The Palestinian Authority has banned pro-[Iraq] rallies, but that ban was widely ignored.

**Athletes, Sports Teams, and GPEs**

Athletes and teams are associated with GPE.GPEs as in example below.

Lance Armstrong of Team United States won the Tour of France 7 times.

However, when a GPE name is used as a team name (as in Boston beat Philly), the entity is marked as a Nickname Metonymy. Please refer to 6.2 for further discussion.

**Political associations**

Political associations hold between people and GPEs. So in Hillary Clinton (D-NY), NY is marked GPE.GPE.

"This is going to be a brutal fight," said Rep. Thomas C. Sawyer D-[Ohio].

Magdy El Sayad National Party-[Beny Sweif] said that the government will support low income people

**3. 4 Locations**

Places defined on a geographical or astronomical basis which are mentioned in a document and do not constitute a political entity give rise to Location entities. These include, for example, the solar system, Mars, the Hudson River, Mt. Everest, and Death Valley.

Places distinguished only by the occurrence of an event at that position ("the scene of the murder", "the site of the rocket launching") are not entities.
3.4.1 Subtypes for Locations
We will further classify Location entities with the following subtypes. Locations that do not fit into the subtypes defined below will not be tagged.

**Address**
A location denoted as a point such as in a postal system or abstract coordinates ("31° S, 22° W"). The name of a location in a postal system is also an address.

```
17 Foad St
```

**Boundary**
A one-dimensional location such as a border between GPE’s or other locations.

```
Lebanese borders
```

**Celestial**
A location which is otherworldly or entire-world-inclusive The rule of thumb is to tag world, earth, globe in addition to all other planets as LOC celestial

```
Scientists say that planet [Mars] may have had lakes after discovery of layers of rocks similar to those on [Earth]
```

**Water-Body**
Bodies of water, natural or artificial (man-made).

```
Combined Egyptian- Saudi air and naval military training began yesterday around the [Red Sea]
```

**Land-Region-natural**
Geologically or ecosystemically designated, non-artificial locations.

```
170 skiers were killed in a fire on a train at a winter resort in [the Alps]
```
**Region-International**
Taggable locations that cross national borders.

ويخطر أيضاً سفر المسؤولين في حركة طالبان حتى حدود رتيبة وزير، الى [الخارج]

**Region-General**
Taggable locations that do not cross national borders.

بعد ظهور اليوم الخميس حدث هجوم مسلح على سيارتهم في [شمال الضفة الغربية]

There was an armed attack Thursday afternoon on cars in [northern **West Bank**]

قتل اثنين عشر شخصا على الأقل في هجوم في [الجزء الذي تسيطر عليه الهند من كشمير]

At least twelve people were killed in an attack in [the Indian **region** of Kashmir]

### 3.4.2 Sub-parts of Locations and GPEs

Portions of GPE entities or Location entities, such as "**the center of the city**", "**the outskirts of the city**", or "**the southern half of New Jersey**" constitute Location entities in their own right. When general locative phrases like “top,” “bottom,” “edge,” “periphery,” “center,” and “middle” are used to pinpoint a portion of a markable location, they are markable locations.

1) The [Sudan the southern] where the southern is an adjective while Sudan is treated as a NOM.LOC This construct occurs rarely in Arabic.

[السودان الجنوبى]

2) South-genetive Sudan where south is a NOM.LOC, while Sudan is a NAM.GPE

[جنوب السودان]

When the text implies “south of the nation of Sudan”, there are two taggable mentions: [South-genetive Sudan] where south is the head (NOM-LOC-Region-General) and [Sudan] is (NAM-GPE-Nation)

3) When the text implies “the southern part of the nation of Sudan”, there is only one mention “Sudan” for the whole extent. It will be Nom LOC

[جنوب السودان]

The construction of 2) and 3) is exactly the same in Arabic, and has to be understood from the text in most of cases. In some other cases it is almost impossible to distinguish between them.
3.4.3 Non-Locations

It is easy to start interpreting all objects as locations. Every physical object implies a location because the space that each physical object occupies is the “location” of that object. In addition, our language is full of location modifiers (which are often prepositional phrases) that pinpoint objects and activities, and even abstract concepts:

"معلقم تحت الكلب"  
"Your coat is under the dog."

"الأربن مختبئ خلف هذه الصخرة"  
"The rabbit is hiding behind that rock."

"لدي فكرة في مخيلتي"  
"I have an idea in my head."

Viewed from a certain angle, “the dog,” “that rock” and “my head” become locations. Very “location-ish” nouns make such an interpretation even more tempting:

"لقد أسقط المفاتيح على الأرض"  
"He dropped the logs on the ground."

"لقد أعاد المصباح إلى مكانه مرة أخرى"  
"He put the lamp back in its place."

However, none of these are taggable location expressions. They do not fall within any of the classes defined above for taggable locations. The annotator must be careful not to fall down this slippery slope.

Do not tag compass points when they serve as adjectives or refer to directions, as in “the ants are heading north” and “they are found as far north as Maine.” Compass points should only be tagged when they refer to sections of a region, as in “the far west.”

These will be annotated as Loc.Reg. International. YES

3.5 Facilities

A facility is a functional, primarily man-made structure. These include buildings and similar facilities designed for human habitation, such as houses, factories, stadiums, office buildings, gymnasiums, prisons, museums, and space stations; objects of similar size designed for storage, such as barns, parking garages and
3.5.1 Subtypes for Facilities
We will further classify Facility entities with the following subtypes. Facility entities which do not fit into the subtypes defined below will not be tagged.

Airport
A facility whose primary use is as an airport.

[New York's La Guardia airport] has been a nightmare this year

Plant
One or more buildings that are used and/or designed solely for industrial purposes: manufacturing, power generation, etc.

Building-or-Grounds
Man-made/-maintained buildings, outdoor spaces, and other such facilities. This includes anything from a tent to a hotel to a ranch to Disneyland.

Hundreds of Palestinians were treated in [Israeli hospitals]

The Israeli army ordered the Palestinians to move out of the [joint communication offices] of the in West Bank and Gaza strip

A Saudi businessman is trying to avoid extradition to the US to be tried on charges of blowing up [the two American embassies]

56 were arrested in the early morning in [their homes]

The relatives of the dead and the wounded attempted to set fire to [the gas station in the city]
Subarea-Facility
Taggable portions of facilities. The threshold of taggability of subarea-facility is the ability of the area to contain a normally proportioned person comfortably. Individual rooms of buildings are considered subarea-facility, but other portions of buildings, such as walls, windows, or doors, are not tagged. 

The first session of the trial was attended by the lawyers in[ the 5th. Charge room in the military court]

Rassam was able to hear the witnesses in Montreal from[ his jail cell in Seattle] via  the closed circuit video

[The hotel lobby]

Path
A facility that allows fluids, energies, persons or vehicles to pass from one location to another.  For example: streets, canals, and bridges.

The Israeli forces prevented the movement of people in [ the main streets]

[Telephone lines] were knocked down……

4 Entity Class
Each taggable entity must be assigned a class that describes the kind of reference the entity makes to something in the world.  The distinction between referential and attributive uses of an NP is given by the following definitions for ACE:

A mention is referential if it (a) introduces a new entity into the discourse or (b) is a definite descriptive term, a name, or an anaphoric expression for a referential mention previously occurred in the discourse.

A mention is attributive if the mention (a) states a property or properties about an entity referenced by another mention within the same sentence – often as an appositive to or part of a predicate on the other mention – or (b) qualifies an entity through immediate modification within the same phrase.
Referential mentions are further divided into generic and non-generic classes. A generic mention refers to a class/kind/species of objects or a typical representative of that class/kind/species and does not point to or pick out any specific individual object(s) of that class/kind/species. So if any property predicates on a generic mention, it means the entire class referred to by the mention has that property, or all/most/any members of that class have the property.

A non-generic referential mention refers to one or more individual member entities of a particular class. The entity or entities can be accounted for by pointing (specific) or cannot be precisely accounted for (underspecified).

Please see Appendix A for the Decision Tree for Entity Class. This tree steps through the process of assigning a class to an entity.

4.1 Negatively Quantified (NEG)

An entity is NEG when it has been quantified such that it refers to the empty set of the type of object mentioned.

[No sensible lawyer] would take that case.

Please note that we do not assign NEG for entities introduced by negated predicates.

They are not [lawyers].

There aren't [any confirmed suspects] yet, but officials say several Middle East groups are expected to be investigated.

4.2 Non-referential/Attributive/Ascriptive (ATR)

An entity is ATR when it is not being used to refer, but rather to attribute some property or attribute to some entity. The titles (3.1.2), the nominal mention in apposition (5.4.1) and postmodifiers (5.1.9) all fall into this class. Note that not all postmodifiers are tagged as ATR. For example, in

[الجيش الإسرائيلي]

Israeli is an ATR; but in

[العرب]
Arab is considered as a SPC. The difference between these two structures is that in “the Army the Israeli”, “the Israeli is a postmodifier of “the army”, while in “summit the Arab” which is equivalent to “the summit of the Arabs” in English, “the Arabs” is a SPC. In the following examples where the title is followed by proper name, the title is annotated as ATR.

[President Al-Asad]  
[Dr. Ali]  
[The chemist Samir ]

If the title is followed by an adjective, it is annotated as SPC, as shown in the following examples:

[The Egyptian president]  
[The smart physician ]  
[The careless student]

4.3 Specific Referential (SPC)
An entity is SPC when the entity being referred to is a particular, unique object (or set of objects), whether or not the author or reader is aware of the name of the entity or its anchor in the (local) real world.

The Palestinian Authority has banned pro-[Iraq] rallies, but that ban has been widely ignored.

[Three Israeli soldiers]  

4.4 Generic Referential (GEN)
An entity is GEN when the entity being referred to is not a particular, unique object (or set of objects). Instead GEN entities refer to a kind or type or class of entity. Notice that the mentions in question are still understood to be referential
in that they point to actual things in the world rather than saying that an object ‘has that property’ or some similar notion. In fact, the subject NP in all the following examples has a generic reading:

In Arabic, an entity must have a definite determiner to be considered generic

[Lawyers] are honest

[Engineers] are creative

[Nurses] take care of the patients

What is common to these examples is the predicate: it’s a kind-level predicate, meaning the predicate describes a property of a kind of entities. Individual- and species-level predicates are also triggers:

[Physicians] like to read new researches

4.5 Under-specified Referential (USP)

We reserve the term underspecified for non-generic non-specific reference. Underspecified references include quantified NP’s in modal, future, conditional, hypothetical, negated, uncertain, question contexts (in all cases the entity/entities referenced cannot be verified, regardless of the amount of “effort”).

[Some Americans] don’t like pizza

Another example of an underspecified entity is a mention of a large number of entities where the actual members of the set are not necessarily identifiable and the number used is an estimate.

[Thousands of people] will attend the conference.

However, if the phrase was in the past tense as in the following example, we will tag it as specific

[Thousands of people] attended the conference.
While we try to define the other four categories as precisely as possible, annotators may still encounter NPs that cannot be classified. In these cases, annotators should make these NPs Underspecified. By partitioning these truly ambiguous cases into the USP category, annotators will be able to make clearer distinctions between the other four categories, thus improving consistency.

5 Mention Types/Mention Levels
For each entity, we record and coreference all mentions of the entity. As mentioned, an entity refers to an object or set of objects in the world. A mention is a reference to an entity. A single entity can have multiple representations, such as proper name, noun phrase, pronoun. If multiple mentions refer to the same entity, we need to coreference them. (In the tool, the way of coreference is to put all mentions of an entity in the same row).

[President [Al Assad]] will give a speech tomorrow. [He] will enumerate all [his] achievements and the problems that [he] still hopes to solve. [The President] will be accompanied by most of the ministers

Mentions will frequently be nested; that is, they will contain mentions of other entities. For example, the phrase

[The president of [Ford]]

is a mention of an entity of type Person, and contains the name "Ford", a mention of an entity of type Organization. It is even possible for a noun phrase to contain an embedded mention of the same entity. For instance,

[[President Mubarak] [الذي] رشح نفسه لفترة رئاسية جديدة]

[[President Mubarak] [who] is running for new presidency period]

Mentions are categorized at several levels. At the top level there are two major types, simple and complex mentions. This top level distinction is motivated by the fact that some mentions have complex syntactic structures that cannot be easily annotated without breaking the syntactic analyses and/or information loss. Subcategorizations of mentions types are syntactically motivated. The following table lists all the types we distinguish for Arabic.
We will tag the mention as PTV only when it contains the preposition “of” as in:

[
khamsa min al-anfalu
]

[Five of the kids]

But for other cases it will be tagged as NOM as in:

[busum min al-anfalu]

[Some of the kids]

5.1 Simple Mentions

Simple mentions are full noun phrases. For each simple mention, we record its full extent and its head.

5.1.1 Mention Extent

The extent of a mention consists of the entire nominal phrase. In case of structures where there is some irresolvable ambiguity as to the attachment of modifiers, the extent annotated should be the maximal extent. In the case of a discontinuous constituent, the extent goes to the end of the constituent, even if that means including tokens that are not part of the constituent. Thus, in

[al-rasid mubarak al-dhi’i’ rasht fursa ranaisa jihihihi]

[President Mubarak[ who is running for another term in office]]

the extent of the mention is the entire phrase:

al-rasid mubarak al-dhi’i’ rasht fursa ranaisa jihihihi

The extent includes all the modifiers of a nominal phrase, including prepositional phrases and relative clauses.

Generally speaking, tokens are broken at white space, and each item of punctuation is treated as a separate character. As a rule, we do not include punctuation such as commas, periods, and quotation marks in the extent of a
mention unless words included within the extent continue on after the punctuation mark.

**Conjoined Mentions that are Modified**
In constructions of conjoined mentions that share the same premodifiers or postmodifiers, each of the conjoined heads is tagged as the head of a single mention. The following examples will yield two tagged mentions.

<table>
<thead>
<tr>
<th>Nom mention</th>
<th>[20 angry men and women]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nom mention</td>
<td>[20 angry men and women]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>NAM</th>
<th>[Bill Clinton and Jimmy Carter who are both former presidents]</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAM</td>
<td>[Bill Clinton and Jimmy Carter who are both former presidents]</td>
</tr>
</tbody>
</table>

**5.1.2 Mention Head**
In addition to the extent of the nominal phrase, the head of the phrase must be marked. In

إعضاء هيئة التدريس

the full mention

/إعضاء هيئة التدريس

and the head is

/إعضاء هيئة التدريس

If the syntactic head of the phrase is a multi-token item, the first right token is marked. If the head is a proper name, however, then the whole extent of the name is considered to be the head. In the following examples, the mention is enclosed in brackets and the head is underlined:

أصبح [أحمد نظيف] رئيس الوزراء الجديد

**5.1.3 Names (NAM)**
Proper nouns and nicknames.
5.1.4 Nominal Constructions (NOM)
A noun quantified with a determiner, a quantifier, or a possessive.

5.1.5 Pronouns (PRO)
Pronouns with the exception of wh-question words and the specifier ‘that’. A pronominal paradigm if Arabic consists of 12 forms: In singular and plural, the 2nd and 3rd persons differentiate gender, while the 1st person does not. In the dual, there is no 1st person, and only a single form for each 2nd and 3rd person. Traditionally, the pronouns are listed in order 3rd, 2nd, 1st.

Personal pronouns

<table>
<thead>
<tr>
<th>dual</th>
<th>Plural</th>
<th>Singular</th>
<th>Person</th>
</tr>
</thead>
<tbody>
<tr>
<td>هما</td>
<td>هم</td>
<td>هو</td>
<td>3rd male</td>
</tr>
<tr>
<td>هما</td>
<td>هن</td>
<td>هي</td>
<td>3rd female</td>
</tr>
<tr>
<td>أنتما</td>
<td>أنتم</td>
<td>أنتم</td>
<td>2nd male</td>
</tr>
<tr>
<td>أنتما</td>
<td>أنتم</td>
<td>أنتم</td>
<td>2nd female</td>
</tr>
<tr>
<td>نحن</td>
<td>أنا</td>
<td>أنا</td>
<td>1st</td>
</tr>
</tbody>
</table>

Notes:
In Arabic the pronoun is normally not used. Instead, there are a lot of enclitic forms of the pronoun which are affixed to nouns (representing genitive case, i.e. possession) and to verbs (representing accusative, i.e. a direct object). However, we do not tag enclitic pronouns as in:

مكتبته حكومتهم جامعتنا
His library          Their government             our university

5.1.6 WH-Question Words and Specifiers (WHQ)
WH-question words.

5.1.7 Headless Mentions (HLS)
Headless mentions are constructions in which the nominal head is not overtly expressed. Although these mentions are technically headless, we will assign as head the quafiers such as the numbers in the following. Note that in Arabic, there are similar structures like the English expression “the toughest”. We treat these as Nominals rather than Headless mentions.

5.1.8 Partitive Constructions (PTV)
Partitive constructions have two elements: the part and the whole. The first element of a partitive construction lacks a head and quantifies over the second element. Just as in Headless mentions, we will tag the right most premodifier of the first element as the head of the partitive construction.

5.1.9 Postmodifier Mentions (NAMPRE, NOMPRE)
Postmodifier mentions (the premodifiers in English) are those mentions which occur in a modifying position before another word(s). It is immaterial whether or not the word being modified is a taggable entity.
In almost all cases, the construction of the Postmodifier mention must be identical to the construction of the mention as it would occur in a NOM, or NAM construction. The only exception to this rule is the transformations that occur on name-mentions of LOC’s and GPE’s in their premodifier positions. Transformations of names or nominal mentions for any other types of entities are not taggable.

---

[The *Israeli*] troops

[The *Lebanese*] publishers

[The *German*] language

<table>
<thead>
<tr>
<th>Taggable</th>
<th>Not Taggable</th>
</tr>
</thead>
<tbody>
<tr>
<td>the [Russian] foreign minister</td>
<td>the [الروسية] وزير الخارجية</td>
</tr>
<tr>
<td>[mountain] regions</td>
<td>[الجبال] المناطق الجبلية</td>
</tr>
<tr>
<td>[government] grants</td>
<td>[الحكومية] المنح الحكومية</td>
</tr>
<tr>
<td>[Alpine] villages</td>
<td>[الألب] قري</td>
</tr>
</tbody>
</table>

---

5.1.10 NAM vs. NOM

Some ambiguities can arise when trying to make a NAM-NOM distinction. It may appear that a NOM is being used to name something, or that a NAM mention may be decomposable into a few NOMs.

A general property of NAMs is that they are defined to pick out one particular entity as a referent. They are unique identifiers, like "Vladimir Putin" or "The United States."

NOMs, on the other hand, define an entire category. They can pick out a referent which belongs to its category, but only after disambiguating it from all
other potential members of its category. If a nominal mention is used as an individual reference in a discourse, the head noun often has to be “individualized” via quantification and/or qualification with determiners, adjectives, relative clauses, etc.

[Al Assad] is a NAM while [The President] is a NOM.

One of the trickiest parts of distinguishing NAM’s and NOM’s is NOM’s modified by NAM’s such that they only have one referent where the NOM refer to an entire category, as:

America’s army

the Egyptian Supreme Court

the Israeli army

In the above examples, army, supreme court are two categories, which makes them two NOM mentions in ACE. But with the GPE modifying the categories, they pick out a specific referent in each category. It is hard to decide whether the whole string should be treated as a NAM or a NOM mention with GPE NAMPRE. Looking up the real name of the entity is not a good option for any entity type, since consistency is more important than accuracy and annotation speed can’t afford to drop. Hence, we adopt the following rules in Arabic:

For such NOMs with embedded GPEs, if the GPE is in the form of noun, and the whole string is an iddafa, treat the whole iddafa as a NAM

America’s army

Ireland’s government

Philadelphia court
If the GPE is in an adjectival form, treat it as a NOM embedded with a GPE, for example:

[Cairo National Bank]

In cases of GPE modified NAMs, we also need to consider whether the GPE is part of the head of the name or a nested GPE entity. For example:

European Automobile Manufacturers' Association  
Chinese Center for Disease Control and Prevention

The rule is to include the GPE as part of the head of the name, even though “Automobile Manufacturers’ Association” or “Center for Disease Control and Prevention” could be stand-alone NAMs. Annotators may often find the two versions of names in the same file.

[European Automobile Manufacturers’ Association]NAM-ORG

[Automobile Manufacturers’ Association]NAM-ORG

[Chinese Center for Disease Control and Prevention]NAM-ORG

[Center for Disease Control and Prevention]NAM-ORG
5.2 Complex Constructions

The purpose of complex constructions is to identify difficult regions where the simple mention extent rules do not apply. We do not identify heads for complex constructions. Within the extent of a complex construction, simple mentions will be annotated. Each of these complex constructions has rules for extent.

5.2.1 Appositive Constructions (APP)

Apposition is a construction which consists of two or more full mentions which refer to (or predicate on) the same entity. The two mentions are placed side by side, with one explains or characterize the other. Superficially this looks similar to a noun modifying another noun. However, they are different in that in an appositive construction, the first NP and the second NP roughly refers to the same entity, whereas this is not the case in a noun-noun compound. In annotation, the APP-mention itself has no head-assignment.

Another structure in Arabic similar to App constructions is the construction of proper names with titles (3.1.2 and 3.1.3), such as “the president Putin” which is not considered as an apposition. The difference is that when the nominal head has modifiers, we treat the whole phrase as an APP; when the nominal head has no modifiers, we tag the nominal as the head of a NOM mention and the proper name as a NAM mention. Below is the annotation for these two structures.

<table>
<thead>
<tr>
<th>Title with proper name</th>
<th>Apposition:</th>
</tr>
</thead>
<tbody>
<tr>
<td>بكين (along the way)</td>
<td>الرئيس [أرفان][ بوتين]</td>
</tr>
<tr>
<td>پوتین (Valmir Putin)</td>
<td>[Russian] [president][ Valmir Putin]</td>
</tr>
<tr>
<td>Russia’s [President Putin] (not an APP in English)</td>
<td>[Secretary General to [ the UN][ Kofi Anan]]</td>
</tr>
<tr>
<td>рега [زمون]</td>
<td>[Russia’s] [President Putin]</td>
</tr>
<tr>
<td>روسيا [روسيا] روسي</td>
<td>[نام מסוים] [الأمم المتحدة] [كوفي عنان]</td>
</tr>
<tr>
<td>Secretary General to [ the UN][ Kofi Anan]</td>
<td></td>
</tr>
</tbody>
</table>
| وراث [POPE] | [Pope [John Paul II] NOM.ATR]
| [John Paul II] NAM. SPC | |
| [John Paul II] NAM. SPC | [John Paul II] NAM. SPC |
| [Secretary General to [ the UN][ Kofi Anan] | [Secretary General to [ the UN][ Kofi Anan]] |
5.2.2 Complex Constructions taking a Relative Clause (ARC)

An ARC-construction is an appositional construction with an adjacent relative clause that is an adjunct to the entire apposition, rather than just to the Nominal mention or Name mention. In ARC-constructions, the component entity mentions and the WHQ mention all are tagged and assigned heads, after which the headless ARC-tag is applied.

6 Nickname Metonymy

Metonymy occurs when a speaker uses a reference to one entity to refer to another entity (or entities) related to it. For example, in the sentence below, \textit{Beijing} is a capital city name that is used as a reference to the Chinese government:

\textit{Beijing will not continue sales of anti-ship missiles to Iran.}

While this phenomenon occurs in many different contexts and to varied degrees, we are only interested in what we are calling “Nickname Metonymy” for the purposes of this stage of the annotation process. This kind of metonymy occurs when the name of one entity is used to refer to another entity. The pure metaphoric mentions as in “he is my sun” is not treated as metonymy, but as attributive nominal. The sections below outline several common examples. This is not an exhaustive list. Any example of this kind of reference should be identified. We will coreference the mention with the entity to which the mention
refers in the context and indicate that this is an example of Nickname Metonymy by selecting the check box.

6.1 Capital City or Government Seat (FAC) Names standing in for Country’s Government

Cases in which the capital city or the building that is the seat of government is used to refer to the nation’s government are marked as classic metonyms.

[Beijing] will not continue sales of anti-ship missiles to Iran.

In this example there are two senses of the word Beijing: the city Beijing and the government of China. We will tag Beijing as the government of China (GPE.ORG) mention and coreference it with the China entity. If there is a later mention of the city of Beijing (for example, Cohen left {the city} this morning), it would be a GPE.LOC mention of the Beijing entity.

Common examples of government seats used to refer to the nation’s government are “The White House” and “The Kremlin”. We will tag these according to the entity to which they refer.

6.2 City name for Sports Team

When the name of a GPE refers to a sports team, the mention is marked for Nickname Metonymy. The mention is coreferenced with the sports team’s entity.

7 Cross-Type Metonymy

Cross-Type Metonymy occurs when more than one aspect of an entity is referenced in a document. For example, entities of type Organization often have a physical entity of type Facility associated with them. These two incarnations of the same entity will be tagged as type Organization when the textual reference is directly referring to the organization and as type Facility when the mention refers to the physical building.

At the entity annotation stage, we will group entities of different types together with a Metonymy relation when they refer to different aspects of the same underlying entity.
The most common Cross-Type Metonymy Link occurs between Organizations and the Facilities they occupy. These two entities are often referred to using the same strings of text.

Examples

In this example, there are two mentions of a hospital. The first mention is referencing the physical building or hospital facility. The second references the organization that runs or administrates the hospital (the sequence is reversed in the Arabic example. The first mention is reflecting the ORG, while the second references the FAC).

We will annotate examples like this as follows.

<table>
<thead>
<tr>
<th>CE</th>
<th>Argument 1</th>
<th>Cross type metonymy</th>
</tr>
</thead>
<tbody>
<tr>
<td>CE1(Arabic)</td>
<td>مستشفى القديس جون مكومب</td>
<td>CE1 (ORG. Med/Sci)</td>
</tr>
<tr>
<td></td>
<td>St. John Macomb Hospital</td>
<td></td>
</tr>
<tr>
<td>CE2 (Arabic)</td>
<td>المستشفى hospital</td>
<td>CE2 (FAC. Buid/Ground)</td>
</tr>
</tbody>
</table>
Appendix:

Some annotation examples:

### FAC.Building

<table>
<thead>
<tr>
<th>Decision</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>FAC.Building</td>
<td>ميناء العقبة</td>
</tr>
<tr>
<td>FAC.Building</td>
<td>ميعتقل جوازاتامو</td>
</tr>
</tbody>
</table>

### FAC.Subarea

<table>
<thead>
<tr>
<th>Decision</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>FAC.Subarea</td>
<td>حدائق الزهور</td>
</tr>
</tbody>
</table>

### GPE.Cluster

<table>
<thead>
<tr>
<th>Decision</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPE.Cluster</td>
<td>دول البلقانcents</td>
</tr>
<tr>
<td>GPE.Cluster</td>
<td>دول الكمنونث</td>
</tr>
<tr>
<td>GPE.Cluster</td>
<td>دول الشرق</td>
</tr>
<tr>
<td>GPE.Cluster</td>
<td>دول الاتحاد الأوروبي</td>
</tr>
<tr>
<td>GPE.Cluster</td>
<td>مجموعة الثمان دول الصناعية الكبرى</td>
</tr>
<tr>
<td>GPE.Cluster</td>
<td>دول الشرق الأوسط</td>
</tr>
<tr>
<td>GPE.Cluster</td>
<td>دول جنوب آسيا</td>
</tr>
<tr>
<td>GPE.Cluster</td>
<td>دول جنوب شرق آسيا</td>
</tr>
<tr>
<td>GPE.Cluster</td>
<td>دول الغرب</td>
</tr>
</tbody>
</table>

### GPE.Continent

<table>
<thead>
<tr>
<th>Decision</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPE.Continent</td>
<td>أفريقيا</td>
</tr>
<tr>
<td>GPE.Continent</td>
<td>أنتراكتيك</td>
</tr>
<tr>
<td>GPE.Continent</td>
<td>أنتراكتيكا</td>
</tr>
<tr>
<td>GPE.Continent</td>
<td>آسيا</td>
</tr>
<tr>
<td>GPE.Continent</td>
<td>الآسيوي</td>
</tr>
<tr>
<td>GPE.Continent</td>
<td>أوروبا</td>
</tr>
<tr>
<td>Decision</td>
<td>Name</td>
</tr>
<tr>
<td>----------</td>
<td>--------------</td>
</tr>
<tr>
<td>GPE.Continent</td>
<td>أمريكا الشمالية</td>
</tr>
</tbody>
</table>

### GPE.County

<table>
<thead>
<tr>
<th>Decision</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPE.County</td>
<td>مقاطعة فيرفاكس</td>
</tr>
</tbody>
</table>

### GPE.Nation

<table>
<thead>
<tr>
<th>Decision</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPE.Nation</td>
<td>أمريكا</td>
</tr>
<tr>
<td>GPE.Nation</td>
<td>أستراليا</td>
</tr>
<tr>
<td>GPE.Nation</td>
<td>محور الشرق</td>
</tr>
<tr>
<td>GPE.Nation</td>
<td>جزر كايمان</td>
</tr>
<tr>
<td>GPE.Nation</td>
<td>الصين</td>
</tr>
<tr>
<td>GPE.Nation</td>
<td>الكونغو</td>
</tr>
<tr>
<td>GPE.Nation</td>
<td>إنجلترا</td>
</tr>
<tr>
<td>GPE.Nation</td>
<td>إيران</td>
</tr>
<tr>
<td>GPE.Nation</td>
<td>العراق</td>
</tr>
<tr>
<td>GPE.Nation</td>
<td>إسرائيل</td>
</tr>
<tr>
<td>GPE.Nation</td>
<td>الكوري</td>
</tr>
<tr>
<td>GPE.Nation</td>
<td>أسبانيا</td>
</tr>
</tbody>
</table>

### GPE.Pop

<table>
<thead>
<tr>
<th>Decision</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPE.Pop</td>
<td>أطلنتا</td>
</tr>
<tr>
<td>GPE.Pop</td>
<td>بغداد</td>
</tr>
<tr>
<td>GPE.Pop</td>
<td>هونج كونج</td>
</tr>
<tr>
<td>GPE.Pop</td>
<td>كاي إيراجو</td>
</tr>
<tr>
<td>GPE.Pop</td>
<td>ماكار</td>
</tr>
</tbody>
</table>

### Gaza

<table>
<thead>
<tr>
<th>Decision</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPE.Pop</td>
<td>غزة</td>
</tr>
</tbody>
</table>
### GPE.Special

<table>
<thead>
<tr>
<th>Decision</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPE.Special</td>
<td>فلسطين</td>
</tr>
<tr>
<td>GPE.Special</td>
<td>تايوان</td>
</tr>
</tbody>
</table>

### GPE.State

<table>
<thead>
<tr>
<th>Decision</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPE.State</td>
<td>قطاع غزة</td>
</tr>
<tr>
<td>GPE.State</td>
<td>أيرلندا الشمالية</td>
</tr>
<tr>
<td>GPE.State</td>
<td>الضفة الغربية</td>
</tr>
</tbody>
</table>

### LOC.Land-Reg

<table>
<thead>
<tr>
<th>Decision</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOC.Land-Reg</td>
<td>جزر الأزور</td>
</tr>
<tr>
<td>LOC.Land-Reg</td>
<td>جزر الكتاري</td>
</tr>
<tr>
<td>LOC.Land-Reg</td>
<td>المندحدر الفناري</td>
</tr>
<tr>
<td>LOC.Land-Reg</td>
<td>ديبجو جراسيا</td>
</tr>
<tr>
<td>LOC.Land-Reg</td>
<td>جزر إيبيرست</td>
</tr>
<tr>
<td>LOC.Land-Reg</td>
<td>الهيمالايا</td>
</tr>
<tr>
<td>LOC.Land-Reg</td>
<td>جزيرة الكورية</td>
</tr>
<tr>
<td>LOC.Land-Reg</td>
<td>جبل إيفيست</td>
</tr>
<tr>
<td>LOC.Land-Reg</td>
<td>جزر نيشانج كان</td>
</tr>
<tr>
<td>LOC.Land-Reg</td>
<td>هضبة سيكاميان</td>
</tr>
<tr>
<td>LOC.Land-Reg</td>
<td>جزيرة أربونان</td>
</tr>
</tbody>
</table>

### LOC.Reg-Gen

<table>
<thead>
<tr>
<th>Decision</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOC.Reg-Gen</td>
<td>وسط جداكارتا</td>
</tr>
<tr>
<td>LOC.Reg-Gen</td>
<td>الساحل الشرقي</td>
</tr>
<tr>
<td>LOC.Reg-Gen</td>
<td>شرق تكساس</td>
</tr>
</tbody>
</table>
### LOC.Reg-Int

<table>
<thead>
<tr>
<th>Decision</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOC.Reg-Int</td>
<td>الشمالي</td>
</tr>
<tr>
<td>LOC.Reg-Int</td>
<td>الشرق الأقصى</td>
</tr>
<tr>
<td>LOC.Reg-Int</td>
<td>آسيا الكبرى</td>
</tr>
<tr>
<td>LOC.Reg-Int</td>
<td>اليابان الصينية</td>
</tr>
<tr>
<td>LOC.Reg-Int</td>
<td>كمبور</td>
</tr>
<tr>
<td>LOC.Reg-Int</td>
<td>مانشو</td>
</tr>
<tr>
<td>LOC.Reg-Int</td>
<td>مانشوريا</td>
</tr>
</tbody>
</table>

### LOC.Water

<table>
<thead>
<tr>
<th>Decision</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOC.Water</td>
<td>البحر العربي</td>
</tr>
<tr>
<td>LOC.Water</td>
<td>الأطلسي</td>
</tr>
<tr>
<td>LOC.Water</td>
<td>خليج شيزابيك</td>
</tr>
<tr>
<td>LOC.Water</td>
<td>نهر الفرات</td>
</tr>
<tr>
<td>LOC.Water</td>
<td>المحيط الهندي</td>
</tr>
<tr>
<td>LOC.Water</td>
<td>نهر خور عبدالله</td>
</tr>
<tr>
<td>LOC.Water</td>
<td>نهر خورس</td>
</tr>
<tr>
<td>LOC.Water</td>
<td>بحيرة جينيف</td>
</tr>
<tr>
<td>LOC.Water</td>
<td>البيلوريات العظمى</td>
</tr>
<tr>
<td>LOC.Water</td>
<td>المحيط الباسفيكي</td>
</tr>
<tr>
<td>LOC.Water</td>
<td>البحر الأحمر</td>
</tr>
<tr>
<td>LOC.Water</td>
<td>خليج سان فرانسيسكو</td>
</tr>
<tr>
<td>LOC.Water</td>
<td>مضيق جبل طارق</td>
</tr>
<tr>
<td>LOC.Water</td>
<td>نهر دجلة</td>
</tr>
<tr>
<td>LOC.Water</td>
<td>البحر الأصفر</td>
</tr>
</tbody>
</table>

### Land Areas Named by Water References

<table>
<thead>
<tr>
<th>Decision</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOC.Water,</td>
<td>الكاريبي</td>
</tr>
<tr>
<td>LOC.Reg-Int</td>
<td></td>
</tr>
</tbody>
</table>
### ORG.Non

<table>
<thead>
<tr>
<th>Decision</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>ORG.Non</td>
<td>حزب الإتحاد</td>
</tr>
</tbody>
</table>

### Palestine

<table>
<thead>
<tr>
<th>Decision</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>PER.Group, GPE.Special</td>
<td>الفلسطينيين</td>
</tr>
<tr>
<td>ORG.Gov</td>
<td>السلطة الفلسطينية</td>
</tr>
<tr>
<td>ORG.Gov</td>
<td>منظمة التحرير الفلسطينية</td>
</tr>
</tbody>
</table>

### NOT TAGGABLE

<table>
<thead>
<tr>
<th>Decision</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOT TAGGABLE</td>
<td>اللاتينية</td>
</tr>
</tbody>
</table>