Annotation Guidelines

-Rachna, Vijay, Saurabh

As part of an active research project , these are the guidelines to follow before giving tags to each of the dialogue utterance

The tagset that we have defined is a shallow level tagset containing the tags

1. **Opening** - The "opening" topic description can be used for all of the little things that people tend to do at the beginning of a conversation:Introduce themselves, greetings, etc., . It's useful for us to know where the opening begins and ends, but the opening doesn't contain particular topics. We don't need opening material to be further segmented.

A	All right,	Opening
A	Amy, how are you doing today?	
B	Fine,	
B	fine.	
A	All right,	
A	I think we know what we're going to speak about.	
B	Yeah.	
A	Uh, I tell you what, I'll start off.	
A	How's that?	
B	Okay,	
B	you go ahead.	Opening
A	Um, I personally think to set a mark with the judicial system and we're talking about criminals, criminal cases that they should bring back hangings on weekends, in public places.	Major Theme
B	Uh-huh.	
B	In public places.	

2. **Major Topic** - This refers to topics whose content largely reflects the conversation structure. The major Topic, is the topic that is talked about the most and the topic that sets the tone of the conversation. There can be

multiple major topic starts and ends as shown in the example below and the annotators are advised to tag accordingly .

 We bracket the major topic using the below two tags.

   a. **Major Topic - start**
   b. **Major Topic - end**

A	Have you all been able to do much as a family these days?	Major Theme
B	Pardon me.	
A	Have you all been able to do much as, uh, as a full family these days, all of you.	
B	Oh, yes,	
B	we've, uh, been, um, we've gone other, you know, all kinds of places,	
B	we go out to dinner,	
A	Yeah.	
B	we take the kids to the park,	
B	and, uh, you know, whenever we, he, my husband has time off or something, we take them to the zoo,	
B	or we, you know, we do everything we can possibly as a family.	

3. **Minor Topic** - This is the part of the conversation that has been side-tracked but is still similar to the major Topic. We don't need Minor Topic to be further segmented. There could be multiple occurrences of different minor topics.
   a. **Minor Topic start**
   b. **Minor Topic end**

A	As a matter of fact, I prefer public television.	
B	Uh-huh.	
A	And, uh, I have, particularly enjoy the English comedies, and the English mysteries.	
B	Yeah,	
B	Yeah,	
B	I watch mysteries too.	
B	Is that what you're referring to?	
A	Yes.	

B	Okay,	
B	yeah.	
B	That is good.	
B	I like, uh,	
B	do you read?	Minor theme-start
A	Vociferously.	
B	Okay,	
B	well that's the reason why I like both of those programs is because they're kind of based on books.	
A	Yes.	
B	And, the plots are more, um, challenging, you know, than the sitcoms of regular T V.	Minor Theme-end
A	Well, unfortunately for us at least here in the United States, we, the only access we have to that of course is public television.	Major Theme - start

In the above example, that had a major topic of television shows, the minor theme is seen to be asked about the reading habits of the person , but it's not an off topic as we see that they relate the plots of books to those of TV shows.

4. **Off Topic** - The dialogues to be tagged in this category are the ones that are side-tracked from the topic of discussion and are very dissimilar to the major Topic. We don't need Off Topic to be further segmented. There could be multiple occurrences of different off topics.

A	so I have concentrated on watching James Bond movies, this week	
B	I used to like him too,	
B	he was my hero.	
B	In fact, I like Sean Connery to this day.	Off topic -start
A	Well, they're showing uh, LIVE AND LET DIE at this moment,	
B	Yeah.	
A	and, it is the first appearance of Roger Moore as double oh seven.	
B	Yeah.	
A	So,	
B	Oh, he can never cut it	

A	No,	
A	no	
A	that's true.	
A	I was just thinking that today, he did, well, in some of the later movies he looks very effeminate, in his early movies.	
B	Yeah,	
B	Yeah,	
B	yeah,	
B	he does.	
A	Whereas, uh, Sean Connery has maintained the, uh, how would you, uh, ruddy look,	
B	Uh-huh,	
B	yeah.	
B	I liked his accent too.	
B	And, he can even be in movies that are not, uh, uh, sexually oriented,	
B	and he still comes off great.	
A	Well, you know he was, when was it, last year or year before last, he was voted the sexiest actor in movies, or something.	
B	He is.	
B	Yeah,	
B	he had my vote double	
A	I mean the man is sixty-two years old.	
B	I don't care	Off topic end
A	But, uh, but otherwise you know we,	Major Theme - end
A	uh, with P B S and so forth.	
A	And I, I particularly enjoy, I don't know whether you've had a chance to watch it on P B S but uh, the series, uh, YES MINISTER?	

Taken from the same conversation as seen in the previous example, here we see a very gradual shift to an off-topic as tagged above. This conversation segment is very focused on the actor in James Bond, which is very dissimilar to the major theme of TV shows.

5. **Closing** - Just as conversations can have openings that aren't about a real topic, they can also be closed in a similar way .

A	But, uh, as, as far as that goes, I, we at least agree on what we enjoy.	Closing
B	Yeah,	
B	that's right.	
A	But, uh, I don't know if there is, there is a time limit on this, so, uh.	
B	Yeah.	
B	Surely we've made it.	
A	I think so too.	
A	Let's just,	
A	it's been very pleasant talking to you,	
B	It was very nice talking to you.	
A	and have a good evening.	
B	You too.	
A	Good night!	
B	Bye-bye.	Closing

- **Topic Descriptions**: - Write a summary (**as a word or a phrase**) of the content between the start and end of major/minor/off topic tags.
  - Opening and closing don't have topic description

			Topic tag	Topic description
B	Real problem the last few days, one of their, uh, young hopeless stars has apparently ruined his knee for this season,		minor topic start	Talking about a player who hurt his knee
B	it was a catcher.			
A	Uh-huh.			
B	And they really don't have an excess of catchers			
B	and it's really kind of, kind of late to, you know, to start, be trying to trade for somebody, especially with the salary cap that the Rangers have.			
A	Yeah.			
B	They're not the highest paying ball team around.			
A	Oh no,			
A	no,			
A	they're not, are they.		minor topic end	
B	But I, you know, I would, I would gladly take the salary of most of the guys, I'll tell you.		off topic start	Talking about the salary
A	Oh, goodness,			
A	I would too,			
A	even the bench warmer, I would.			
B	Yeah,			
B	that's right.		off topic end	

How to Annotate -
- The annotators are advised to first **go through the conversation at hand before starting the annotation process**
- After the first read, the annotator now has a general understanding of the topic flow of the conversation
- Start the annotation process with the above defined tags in your second read.

General Queries -

- In case of doubts about tagging a chunk of conversation as either the minor topic or off -topic, Please go through the topic flow of the conversation carefully choose the tag that you feel that defines it more effectively.
- A conversation **doesn't necessarily have to start with an Opening and End with a Closing tag**. There are chances of a conversation starting or ending with Major/Minor/Off topic tags.
- Minor topic start is **always preceded or accompanied** by a major topic start and minor topic end is **always succeeded or accompanied** by a major topic end

Speaker	Conversation	Topic tag	Topic tag	Topic description
B	Real problem the last few days, one of their, uh, young hopeless stars has apparently ruined his knee for this season,	minor topic st ▾	major topic st ▾	Talking about a player who hurt his knee
B	it was a catcher.	▾	▾	
A	Uh-huh.	▾	▾	
B	And they really don't have an excess of catchers	▾	▾	
B	and it's really kind of, kind of late to, you know, to start, be trying to trade for somebody, especially with the salary cap that the Rangers have.	▾	▾	

- The Structure of the conversation should follow the rules as follows:
  - Major topic can have minor and off-topics but Minor topics or Off topics cannot have any layers i.e., they can't have any embedded Minor/Off topic tags
  - Example
    - Major topic start -> Minor topic start ->Minor topic end -> Off topic start->Off topic end ->Major topic end
    - Major topic start -> Minor topic start -> Off topic start->Off topic end ->Minor topic end ->Major topic end

      We see that the second structure disobeys our rules for maintaining the structure of topics in a conversation and thus is wrong. The above wrong structure could be corrected in this manner.

- Major topic start-> Minor topic start -> Minor topic end -> Off topic start->Off topic end ->Minor topic start -> Minor topic end ->Major topic end