Extracting Space Situational Awareness Events from News Text

Zhengnan Xie, Alice Saebom Kwak, Enfa George, Laura W. Dozal, Hoang Van,
Moriba Jah, Roberto Furfaro, Peter Jansen

The University of Arizona

The University of Texas at Austin

Motivation: NLP for Space Situational Awareness

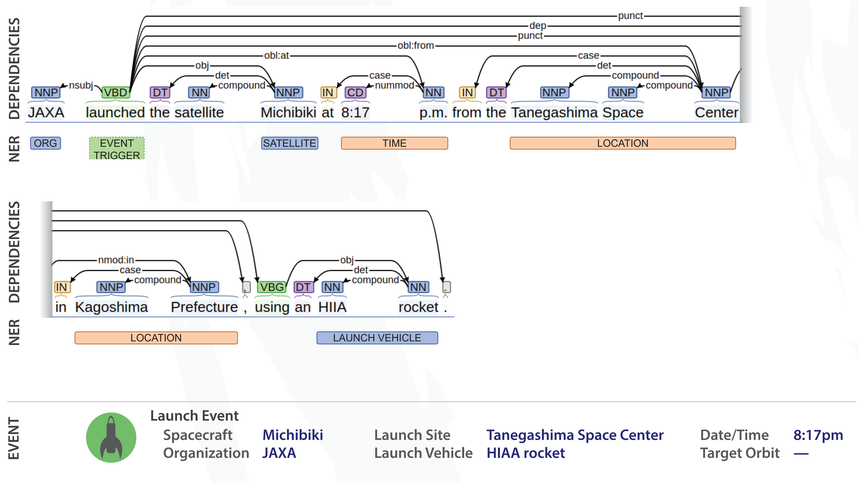
Space situational awareness (SSA) is the task of monitoring the satellites and other space assets of different geopolitical entities.

SSA is normally approached as a **physical measurement** problem, with telescopes and other instruments pointed at the sky.

This work: Can we read text (e.g. news articles) posted on the web to help complement the information gathered using physical measurements?

Approach: SSA is a very low resource domain -- satellites are rarely launched, and events rarely happen to them. Here, we generate a corpus of SSA events using boostrapping, then model this as a span-based information extraction task.

Extraction Rules



Space Situational Awareness News Corpus

48.5k News Articles

Sentence Identification Using Dependency Rules

```
# Launch Vehicle with Satellite
  - name: launchRule
    type: event
    pattern: |
      trigger = [(lemma=/(launch|send|place)/ & tag=/V.*/)]
      org = >nsubj ${org}
      org? = >> ${org}
      vehicle = >dobj >amod? >compound? ${vehicle} ${numberGeneric}? /rocket|launcher/?
      satellite = >dobj? >nmod_with ${satellite} ${numberGeneric}? /satellites?/?
      launchSite? = >/nmod_from|advcl_from/ [entity=/(ORGANIZATION|LOCATION)/]+
      targetOrbit? = >/nmod_into|nmod_in|nmod_at/ ${chunk} orbit
```

2.8M Sentences

Human Annotated Events

Mars Express , **Europe** ' s entry into the ongoing and slowly expanding exploration of the Red Planet from Earth as precursors to later missions by human explorers , was launched on **June 2 , 2003** , from the **Baikonur launch site** by a **Russian** **Soyuz / Fregat** rocket .

1,787 Sentences

Extracted Events

Launch Event

Spacecraft	Starlink	Date	Sept 3 2020
Launch Vehicle	Falcon 9	Organization	SpaceX (US)
Launch Site	Kennedy Space Center		
Target Orbit	Low Earth Orbit (LEO)		

Failure Event

Failure Type	Communications	Date	Mar 27 2016
Spacecraft	Hitomi Telescope	Organization	JAXA (Japan)
Launch Vehicle	—		

Decommissioning Event

Spacecraft	Tiangong-1 Space Station	Date	—
		Organization	China

Corpus Statistics

	Training			Development			Test		
SSA Event	Sents.	Tokens *(w/tag)*	Tokens *(total)*	Sents.	Tokens *(w/tag)*	Tokens *(total)*	Sents.	Tokens *(w/tag)*	Tokens *(total)*
Launch	537	5,646	25,855	350	2,890	13,357	350	3,059	12,747
Failure	310	2,748	12,905	63	580	2,656	63	449	2,043
Decommissioning	81	396	3,064	17	75	487	17	68	504

Table 1: Overall statistics for the space situational awareness event corpus across the three events. Sentences represents the number of sentences within a given set. Tagged tokens represents the number of *BIO-tagged* tokens that have a either a *Beginning (B)* or *Inside (I)* tag.

Results

We frame this as a span-labelling task, and report performance using an off-the-shelf BERT-based NER labeller.

	Development				Test				
Event and Slot	Pr	Re	F1	N	Pr	Re	F1	N	Example
Launch Event				*350*				*349*	
SatelliteName	83	89	86	540	85	89	87	528	Hubble Space Telescope
LaunchVehicle	89	89	89	141	86	92	89	170	Space Shuttle STS-32
LaunchSite	86	97	91	88	89	88	88	137	Cape Canaveral
TargetOrbit	62	84	71	25	89	82	74	38	Low-Earth Orbit
Failure Event				*63*				*63*	
SatelliteName	76	87	81	70	69	87	77	53	Telkom-3
LaunchVehicle	80	84	82	38	88	94	91	31	Proton-M
FailureType	41	55	47	65	48	66	56	50	Launch or power failure
Decommissioning Event				*17*				*17*	
SatelliteName	61	85	71	20	44	67	53	18	NOAA-19
Generic Slots Across Events									
Organization	66	81	73	349	71	82	76	375	NASA
Date	83	91	87	434	84	88	86	390	April 4, 1990

Results using an off-the-shelf BERT-based span labller: https://github.com/kamalkraj/BERT-NER

Corpus Download Link

Space Situational Awareness Corpus available at:
www.github.com/cognitiveailab/ssa-corpus