

UNDP Crisis Response Unit brownbag discussion

## Forecasting Civil Unrest Using Ethnography and Artificial Intelligence

with

Prof Eduardo Albrecht  
(Mercy College, New York)

Prof Melih Kandemir joining remotely  
(Data Science, Ozyegin University)

Dr Stephen Leonard – joining remotely  
(Linguistic Anthropology, Oxford University)

Date: Thursday, 18 April 2017

Time: 1.00 – 2.00pm

Place: DC-1 Building – 20th floor, Julia Tafts Conference Room

**Early warning software** is being used in a wide variety of fields.

Basic concept: Collects open source data from the web to make predictions.

**Health forum example.** Yesha, R. and A. Gangopadhyay, "Social Media Analytics for Behavioral Health." *International Journal of Emergency Mental Health and Human Resilience*.

Possible to predict likelihood of:

- Anxiety
- Depression
- Panic
- Suicide

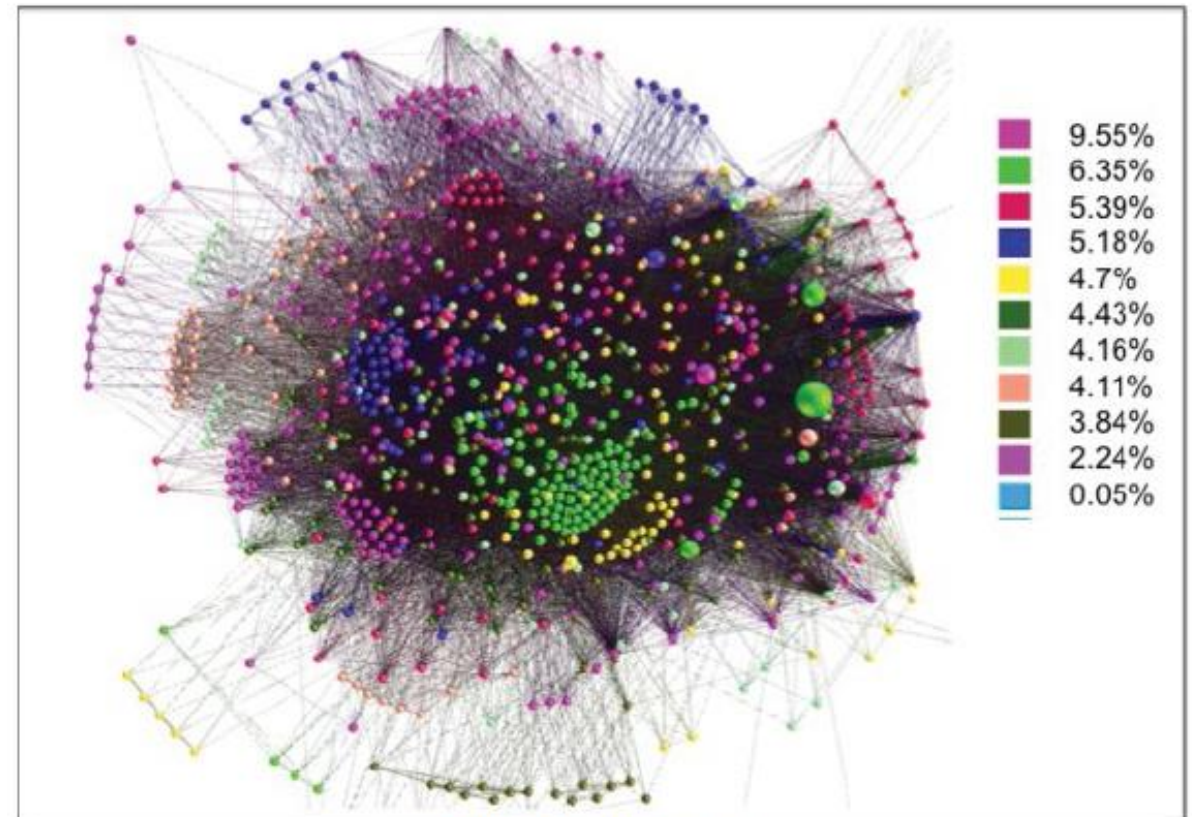


Figure 1. Partitioned Graph

Software tracks and predicts web activity, which is:

- Measurable – providing evidence-based risk assessment.
- Massive – large quantities of information support decision making.

Produces predictions that are:

- Specific, time bound, and verifiable
- Expressed as a percentage of possibility
- Based on past, empirical evidence

(i.e. forum user X has 90% chance of posting suicidal warning sign within 2 weeks)

In my prior consultancy work with private organizations (energy sector), I have found the **key advantages** of using this software are:

- Cover multiple topics and geographical areas with limited resources.
- Integrate empirically based predictions into analysis writing.
- Monitor both anticipated and unanticipated threats to operations.\*

---

\*Low computing costs allow monitoring of potentially unlimited topics.

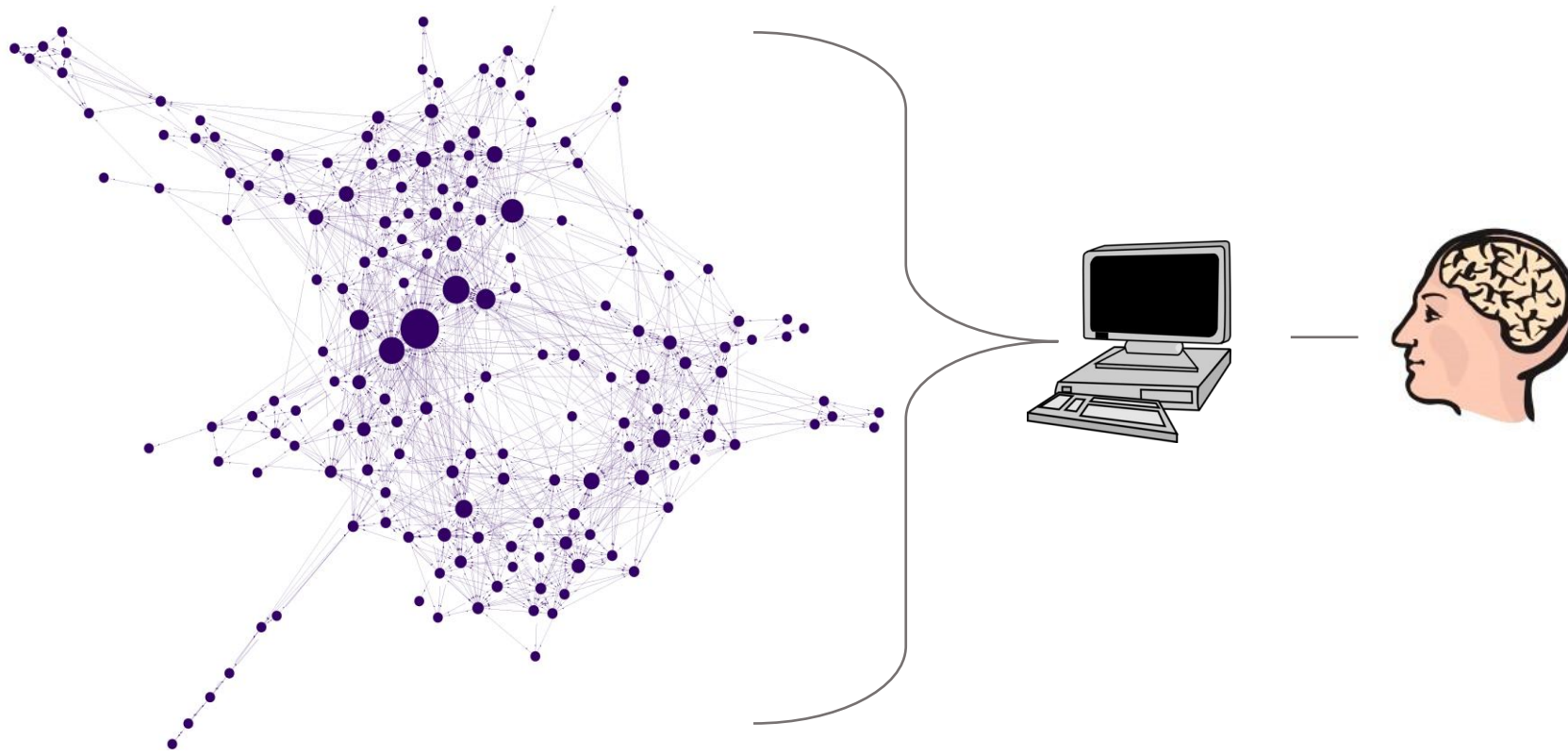
## Allows organizations to:

- Maintain a forward looking mindset.
- Adapt to fast-evolving socio-political contexts.
- Retool for crisis prevention, governance, or recovery scenarios.
- Protect progress toward sustainable goals in politically challenging environments.

Ethnographic Edge uses early warning software to monitor and predict political tensions in **196 geographical areas** worldwide.

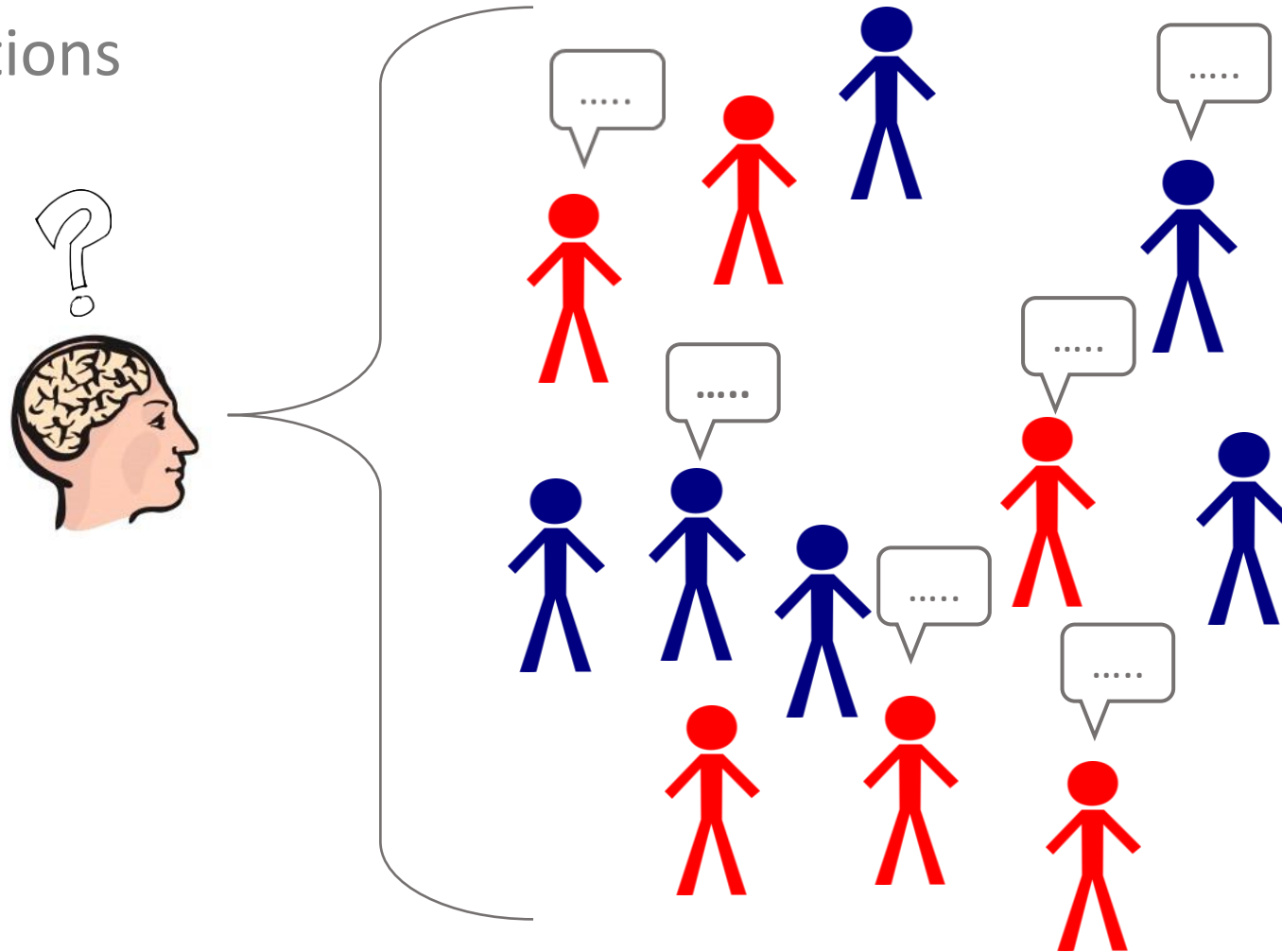
But we combine this methodology with anthropological techniques.

First, **artificial intelligence** is used to analyze large quantities of online data.

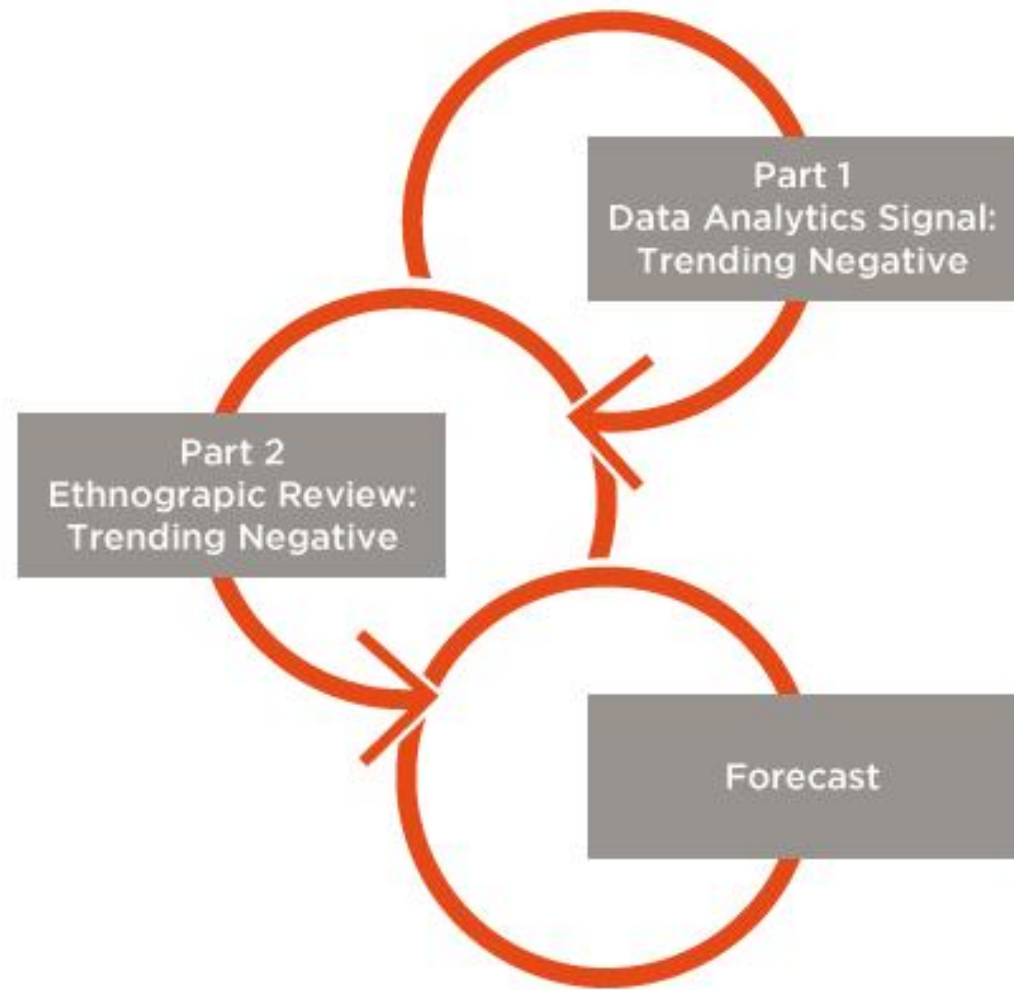
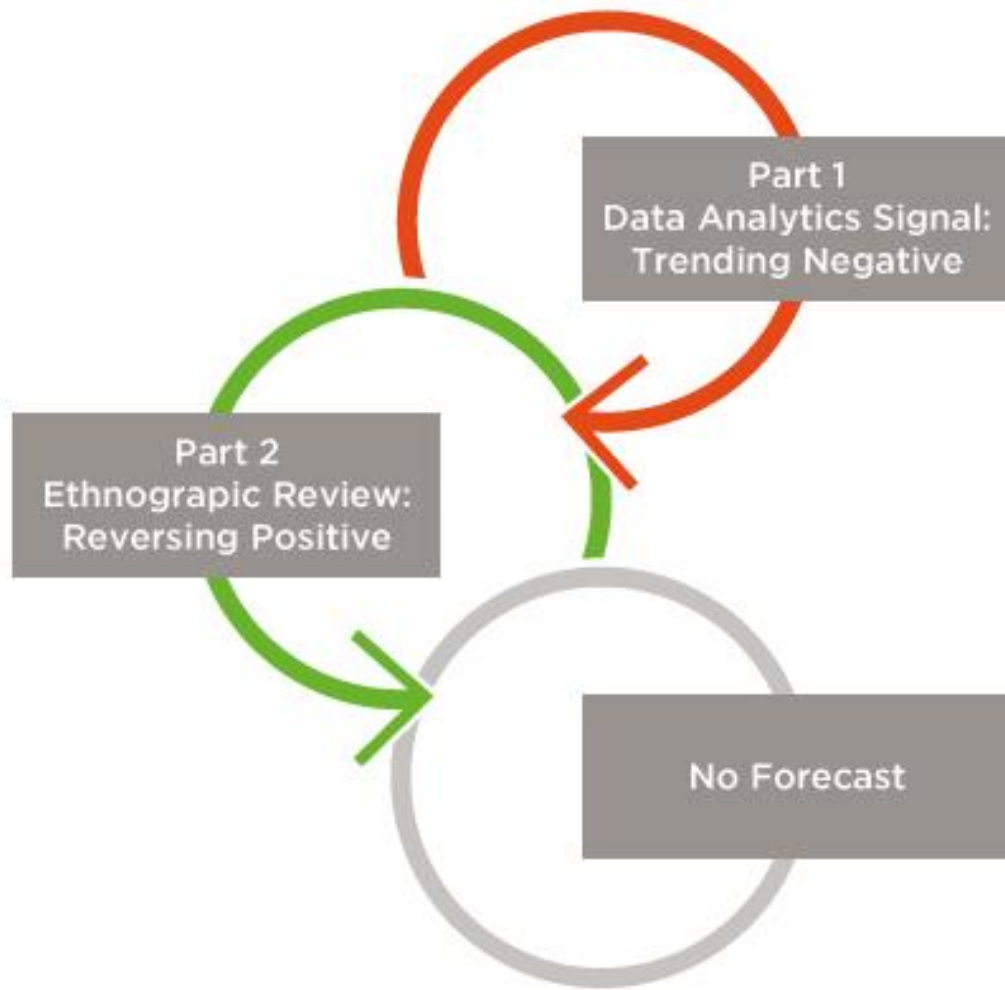


Second, **ethnography** is used to take into account people's perspectives, to:

- recalibrate the software
- refine predictions





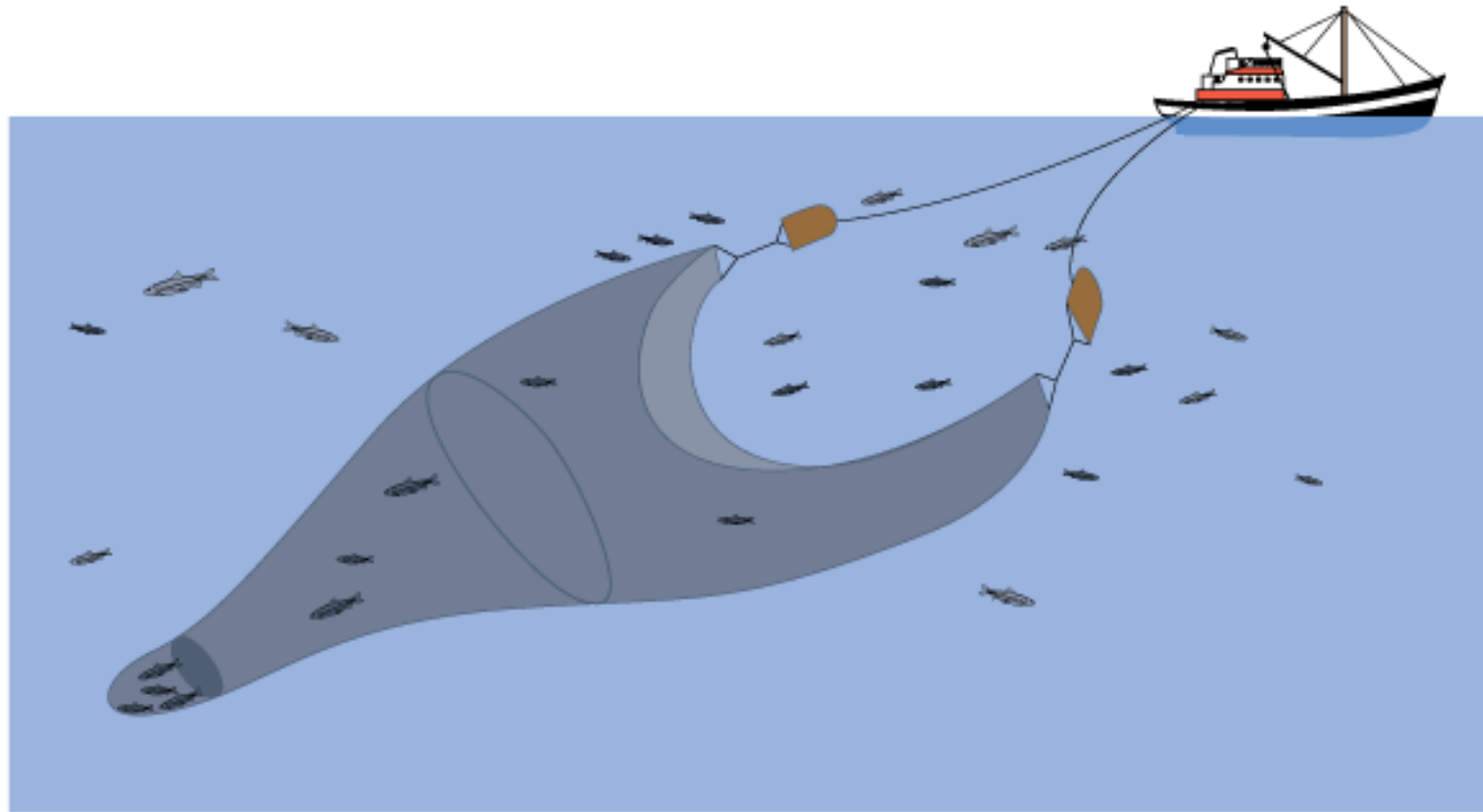


Ethnographic Edge has been collaborating with Recorded Future and using their **Natural Language Processing** (NLP) technologies since 2011.

Recorded Future is an internet technology company founded in 2009 specializing in organizing the entire web for analysis.



This company scans the web using algorithms called **trawlers**. Trawlers search for publications concerning a particular topic.



Information in the publications is then broken down, ranking such measures as **positive/negative sentiment** concerning a topic over time.

For a simplified example, if a **blogger** writes on **May 18**,

**“Anti-government protests expected next Tuesday in Bangkok”**

Negative Sentiment: 2

Event date: May 23

Place: Bangkok

Media type: Blog

Publication date: May 18

Ethnographic Edge App collects this data from thousands of sources (5 years).

Topic: South Sudan														
A	BQU	BQV	BQW	BQX	BQY	BQZ	BRA	BRB	BRC	BRD	BRE	BRF	BRG	
1	Topic: South Sudan	Apr 18 2017	Apr 19 2017	Apr 20 2017	Apr 21 2017	Apr 22 2017	Apr 23 2017	Apr 24 2017	Apr 25 2017	Apr 26 2017	Apr 27 2017	Apr 28 2017	Apr 29 2017	Apr 30 2017
2	Blog Predictive Instances	3658	3355	2989	2243	511	771	2312	2056	1922	2485	2289	1219	1480
3	Blog Predictive Positive	1636	1335	822	807	173	340	932	671	425	687	748	510	515
4	Blog Predictive Negative	218	150	396	319	60	30	85	204	261	136	54	38	103
5	Blog Predictive Neutral	1804	1870	1771	1117	278	401	1295	1181	1236	1662	1487	671	862
6	Government Predictive Instances	66	41	40	64	1	6	45	93	74	7	2	2	4
7	Government Predictive Positive	5	6	3	24	1	0	1	66	16	1	1	0	1
8	Government Predictive Negative	1	2	2	0	0	1	0	0	3	0	0	0	0
9	Government Predictive Neutral	60	33	35	40	0	5	44	27	55	6	1	2	3
10	Mainstream News Predictive Instances	6514	6887	7176	5818	2364	2255	8384	5600	5886	5714	5316	2370	2079
11	Mainstream News Predictive Positive	3168	3336	3488	2849	1065	1012	3343	2993	3199	3193	2336	1234	1093
12	Mainstream News Predictive Negative	470	439	560	382	179	246	569	438	417	464	515	260	253
13	Mainstream News Predictive Neutral	2876	3112	3128	2587	1120	997	4472	2169	2270	2057	2465	876	733
14	Social Media Predictive Instances	208	338	242	264	166	116	140	216	202	166	170	137	144
15	Social Media Predictive Positive	28	61	41	68	19	18	28	66	36	21	32	31	26
16	Social Media Predictive Negative	4	4	18	32	19	6	6	20	27	20	9	20	13
17	Social Media Predictive Neutral	176	273	183	164	128	92	106	130	139	125	129	86	105
18	Blog Not Predictive Instances	9999	9999	9999	9999	9999	9999	9999	9999	9999	9999	9999	9999	9999
19	Blog Not Predictive Positive	4318	4623	4301	4165	3857	4413	4645	4581	4393	4646	4520	4468	4168
20	Blog Not Predictive Negative	1505	1290	1251	1204	1219	1081	1197	1177	1308	1189	999	1036	1034
21	Blog Not Predictive Neutral	4176	4086	4447	4630	4923	4505	4157	4241	4298	4164	4480	4495	4797
22	Government Not Predictive Instances	3570	4725	3041	3437	407	330	2349	4150	2344	366	468	245	249
23	Government Not Predictive Positive	208	757	294	354	40	39	234	461	385	37	85	61	23
24	Government Not Predictive Negative	33	205	49	32	9	14	48	40	43	15	25	10	4
25	Government Not Predictive Neutral	3329	3763	2698	3051	358	277	2067	3649	1916	314	358	174	222
26	Mainstream News Not Predictive Instances	9999	9999	9999	9999	9999	9999	9999	9999	9999	9999	9999	9999	9999
27	Mainstream News Not Predictive Positive	4965	4922	4616	4765	4986	5143	5086	5214	4735	5063	4615	4261	4234
28	Mainstream News Not Predictive Negative	932	918	1227	1165	1058	909	1031	862	1081	906	1225	1240	1598
29	Mainstream News Not Predictive Neutral	4102	4159	4156	4069	3955	3947	3882	3923	4183	4030	4159	4498	4167
30	Social Media Not Predictive Instances	9999	9999	9999	9999	9999	7196	8965	9999	9999	9999	9999	8116	9999
31	Social Media Not Predictive Positive	1605	1495	1420	1538	1305	1278	1784	1603	1933	1800	2175	1224	1460
32	Social Media Not Predictive Negative	1044	926	1113	1408	1306	1092	1294	1973	1506	1993	1686	1523	1519
33	Social Media Not Predictive Neutral	7350	7578	7466	7053	7388	4826	5887	6423	6560	6206	6138	5369	7020

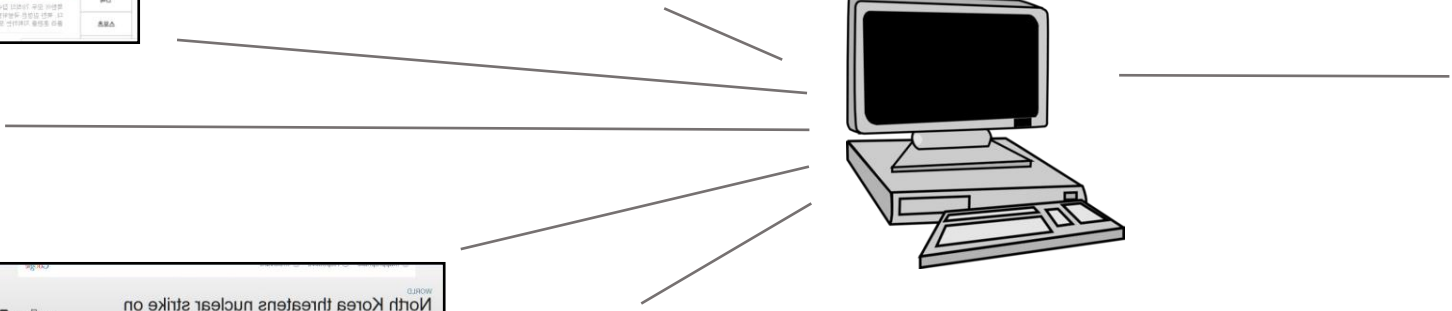
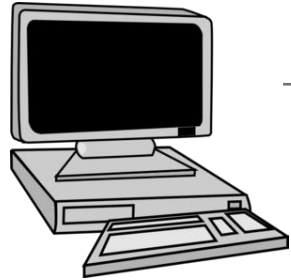
The app interface allows us to **quickly alter** the media types and topics analyzed.

The screenshot displays the 'ETHNOGRAPHIC EDGE' application interface. At the top, there is a navigation bar with the logo and the text 'ETHNOGRAPHIC EDGE'. Below this is a secondary navigation bar with various menu items: Dashboard, Crisis Analysis, Asset Analysis, Crisis News, Crisis Predict, RF Assets, Yahoo Finance, Google Trends, Settings, Users, and Logs. The main content area is titled 'Settings - Entity Exceptions: Recorded Future Crises - News'. A modal dialog box titled 'Add Entity Exception' is open in the center, featuring four dropdown menus for 'Entity', 'Topic', 'Media', and 'Predictive', each with a 'Select' placeholder. The dialog also includes 'Save' and 'Close' buttons at the bottom. In the background, a table lists various entities and their associated topics and predictive statuses. The table has columns for Entity, Topic, and Predictive. The visible rows are: Europe (Business/Finance, Predictive), France (Business/Finance, Predictive), Germany (Business/Finance, Predictive), and Italy (Business/Finance, Not Predictive). Each row has an 'Actions' column with 'Edit', 'Delete', and 'Clear History Data' buttons.

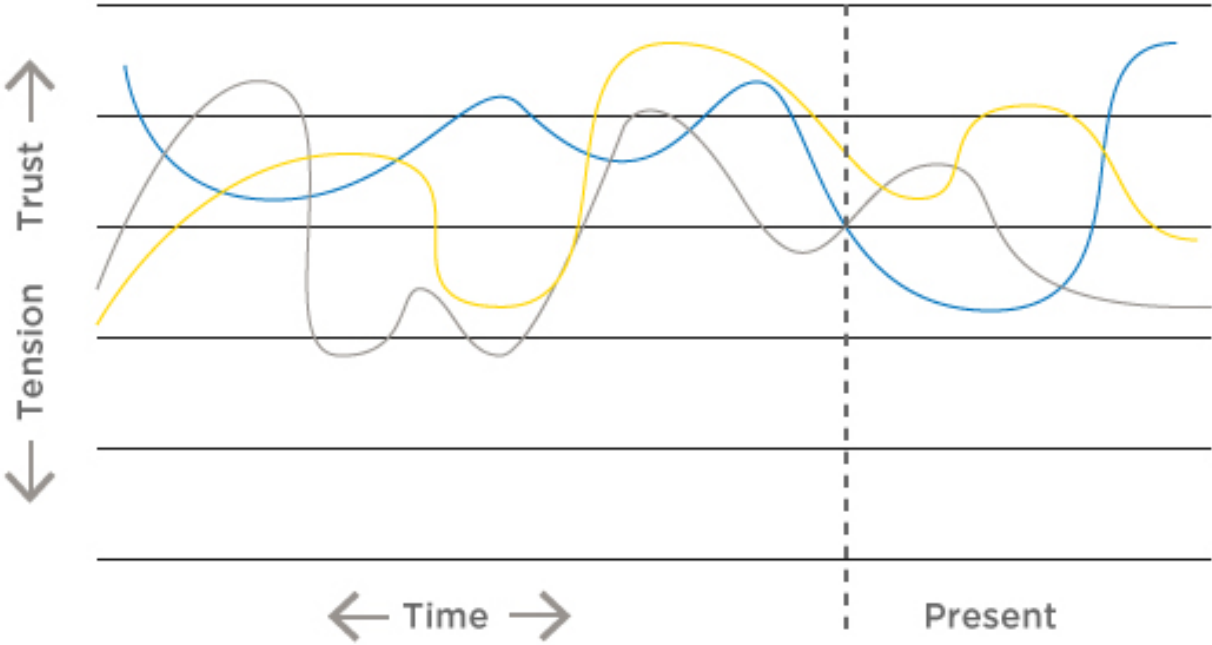
Entity	Topic	Predictive	Actions
Europe	Business/Finance	Predictive	Edit, Delete, Clear History Data
France	Business/Finance	Predictive	Edit, Delete, Clear History Data
Germany	Business/Finance	Predictive	Edit, Delete, Clear History Data
Italy	Business/Finance	Not Predictive	Edit, Delete

# And choose between 12 different languages:

Arabic, Chinese, English, French, German, Italian, Japanese, Farsi, Portuguese, Russian, Spanish, and Swedish



Machine learning software is then used to analyze the data and make predictions.



Recurring combinations are identified and categorized according to their effect on future events.

Present day data is collected and monitored for the same combinations.

Combinations are categorized as either:  
Trending Positive  
Trending Negative  
Reversing Positive  
Reversing Negative





Predictions are made concerning sentiment in a media type. **For example:**  
 Negative mainstream media coverage will increase next week, indicating tensions.

	A	BRB	BRC	BRD	BRE	BRF	BRG
1	Topic: South Sudan	Apr 25 2017	Apr 26 2017	Apr 27 2017	Apr 28 2017	Apr 29 2017	Apr 30 2017
2	Blog Predictive Instances	2056	1922	2485	2289	1219	1480
3	Blog Predictive Positive	671	425	687	748	510	515
4	Blog Predictive Negative	204	261	136	54	38	103
5	Blog Predictive Neutral	1181	1236	1662	1487	671	862
6	Government Predictive Instances	93	74	7	2	2	4
7	Government Predictive Positive	66	16	1	1	0	1
8	Government Predictive Negative	0	3	0	0	0	0
9	Government Predictive Neutral	27	55	6	1	2	3
10	Mainstream News Predictive Instances	5600	5886	5714	5316	2370	2079
11	Mainstream News Predictive Positive	2993	3199	3193	2336	1234	1093
12	Mainstream News Predictive Negative	438	417	464	515	260	253
13	Mainstream News Predictive Neutral	2169	2270	2057	2465	876	733
14	Social Media Predictive Instances	216	202	166	170	137	144
15	Social Media Predictive Positive	66	36	21	32	31	26
16	Social Media Predictive Negative	20	27	20	9	20	13
17	Social Media Predictive Neutral	130	139	125	129	86	105
18	Blog Not Predictive Instances	9999	9999	9999	9999	9999	9999
19	Blog Not Predictive Positive	4581	4393	4646	4520	4468	4168
20	Blog Not Predictive Negative	1177	1308	1189	999	1036	1034
21	Blog Not Predictive Neutral	4241	4298	4164	4480	4495	4797
22	Government Not Predictive Instances	4150	2344	366	468	245	249
23	Government Not Predictive Positive	461	385	37	85	61	23
24	Government Not Predictive Negative	40	43	15	25	10	4
25	Government Not Predictive Neutral	3649	1916	314	358	174	222
26	Mainstream News Not Predictive Instances	9999	9999	9999	9999	9999	9999
27	Mainstream News Not Predictive Positive	5214	4735	5063	4615	4261	4234
28	Mainstream News Not Predictive Negative	862	1081	906	1225	1240	1598
29	Mainstream News Not Predictive Neutral	3923	4183	4030	4159	4498	4167
30	Social Media Not Predictive Instances	9999	9999	9999	9999	8116	9999
31	Social Media Not Predictive Positive	1603	1933	1800	2175	1224	1460
32	Social Media Not Predictive Negative	1973	1506	1993	1686	1523	1519
33	Social Media Not Predictive Neutral	6423	6560	6206	6138	5369	7020

This aspect is addressed by:

**Melih Kandemir**, Project's Data Scientist

Assistant Professor, Ozyegin University, Istanbul

Postdoc, University of Heidelberg

PhD, Aalto University

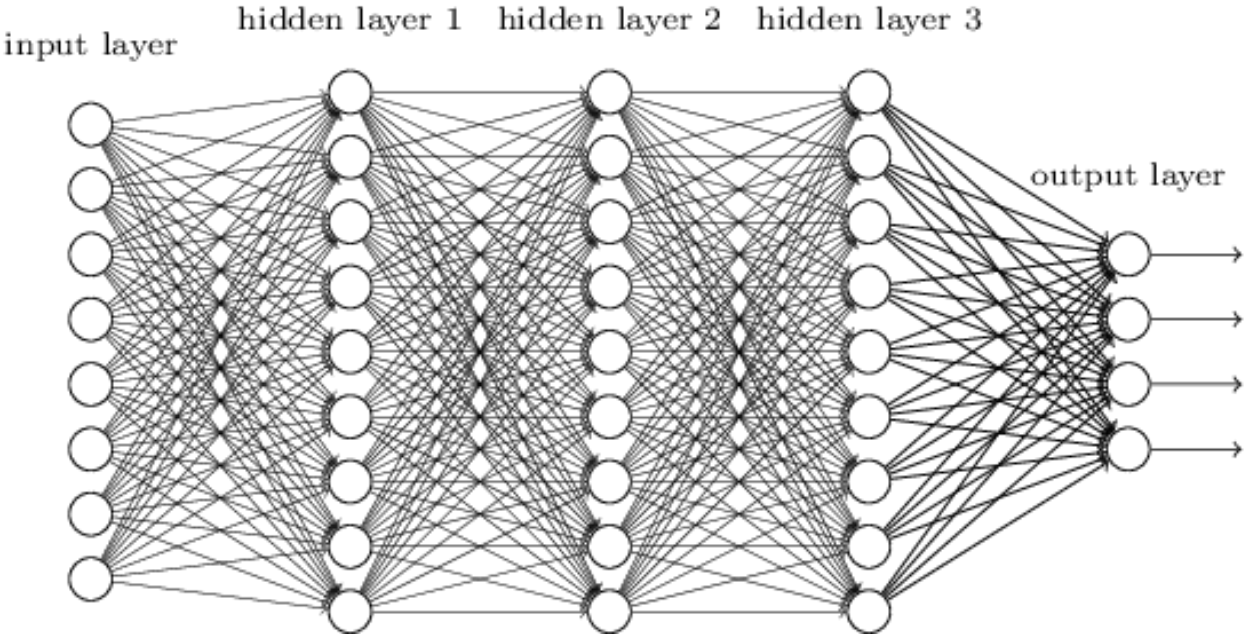
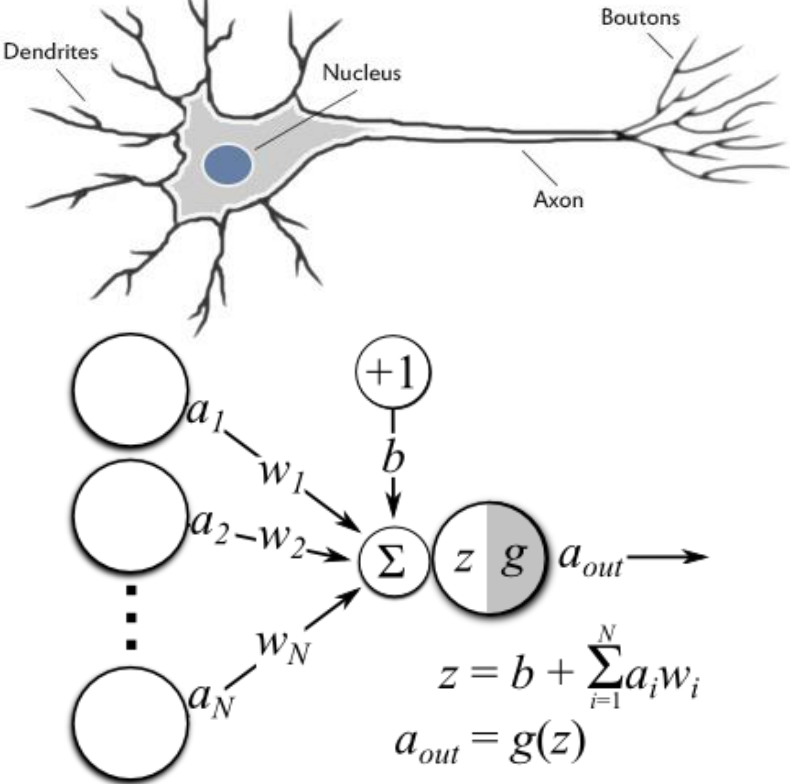
Research activity focuses on novel Bayesian inference techniques for deep learning, and their applications to computer vision, time series analysis, and medical image analysis problems.



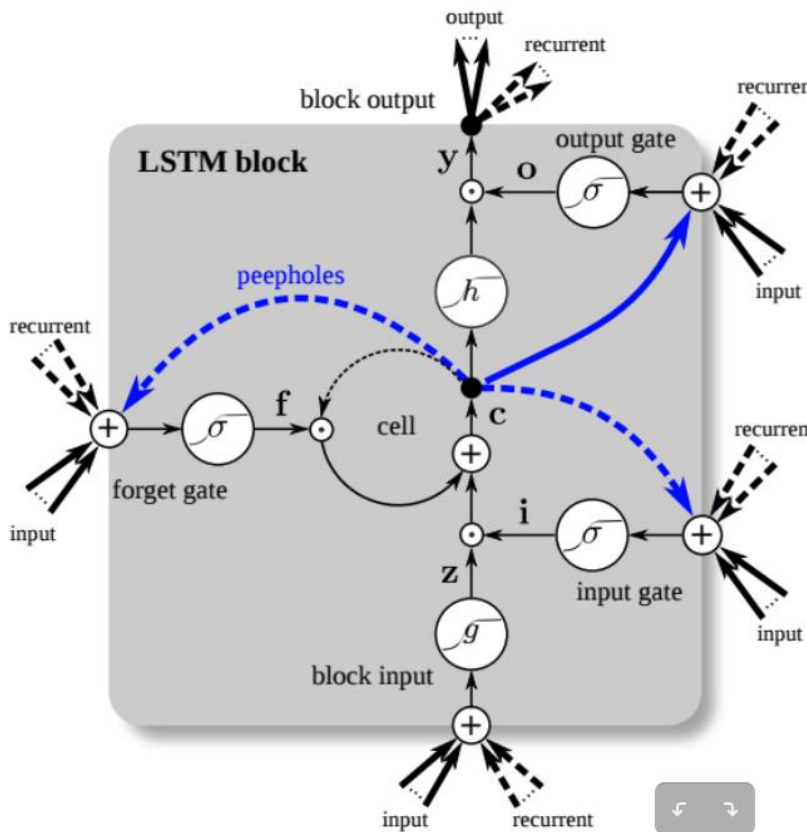
We perform machine learning using a state-of-the-art approach called **Deep Learning**. We train an artificial neural network on past events to make accurate predictions for the future.



An artificial neural network is a mathematical model consisting of a network of simple information processing units, called **neurons**, just as the human brain. We build deep learning models by creating many neurons and linking them to construct layered architectures (i.e. a simulated brain).



Our predictor is a neural network architecture, called the **Long Short-Term Memory (LSTM)**, tailored specifically for time series analysis. We implement our LSTM under TensorFlow, an effective artificial intelligence platform supported by Google.



Google



Next, we use a linguistic **anthropological framework** to critically assess the online data and the outcomes of the model, and to recalibrate the data and the model.

A screenshot of Java code displayed on a dark background with light blue text. A magnifying glass is positioned over the code, focusing on the `decodeMessage` method. The code includes several methods: `extractMessage`, `decodeMessage`, and another `extractMessage`. The magnified section shows the `decodeMessage` method signature and its body, which includes a loop for `MAX_RES` and a call to `extractMessage`.

```
for (int j = 0; j < loc; j++) res[j] = buf[j];
return res;

public void ... (int[] res) {
    for (int i = 0; i < res.length; i++) {
        res[i] = checkRes[i];
    }
}

public int[] decodeMessage(int[] res) {
    for (int i = 0; i < MAX_RES; i++) buf[i] = 0;
    int i = 0;
    while (i < res.length) {
        ... (1000 + i);
        ... buf[loc];
        ... RES_LEN);
    }
}

return null;

public int[] extractMessage(int[] res) {
    for (int i = 0; i < MAX_RES_LEN; i++) buf[i] = 0;
    int loc = 0, i = 0;
    while (i < res.length) {
```

This aspect is being addressed with:

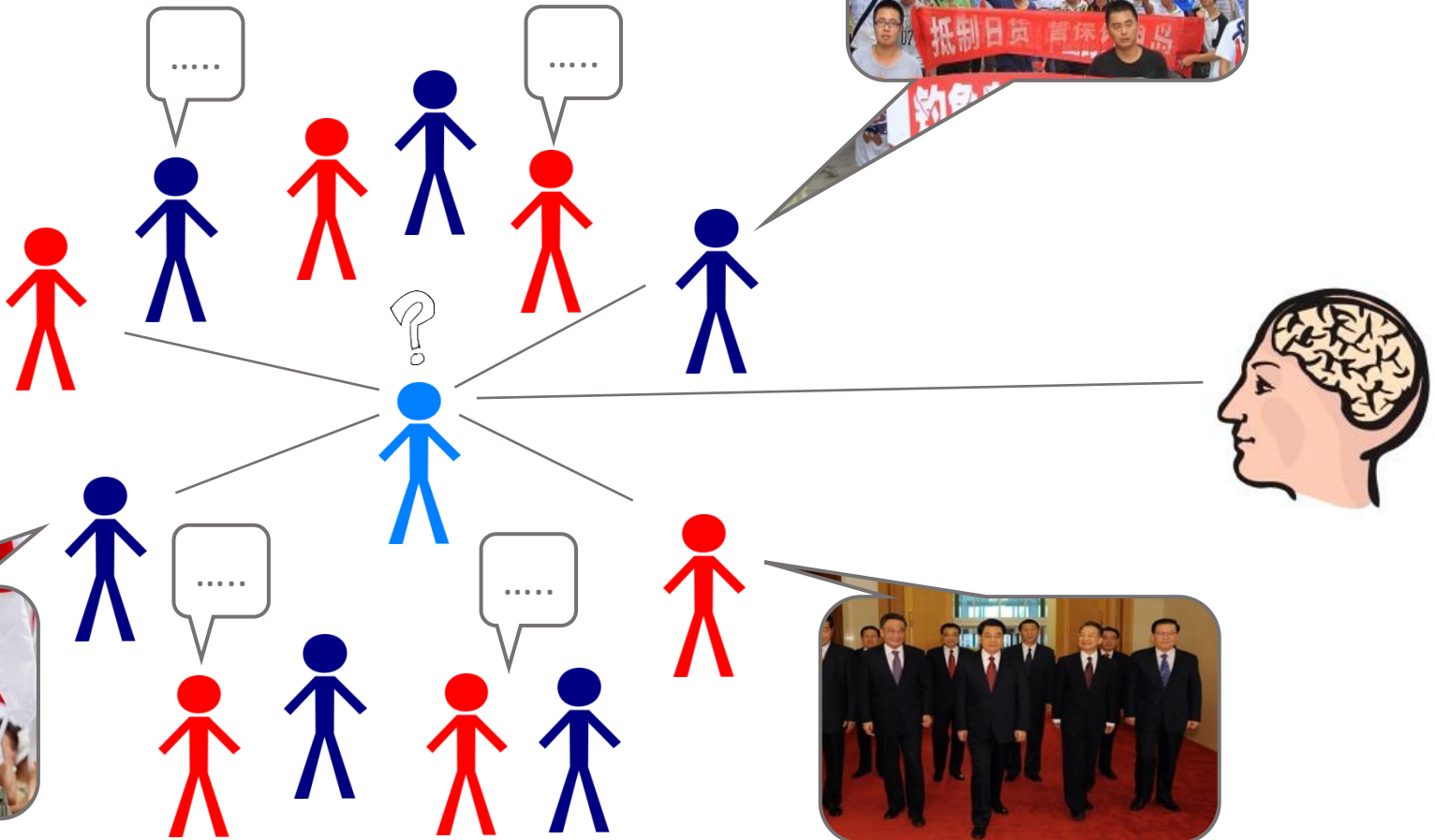
**Stephen Pax Leonard**, Linguistic Anthropologist

Fellow, Exeter College, University of Oxford  
Research Fellow, Trinity Hall, Cambridge  
D.Phil, University of Oxford

Publishes on arctic language and identity, questions relating to existential anthropology, and the relationship between language, ideology, and power.



Lastly, through a network of **160 informants** worldwide we can collect local knowledge concerning political developments.





Ethnographic Edge has been collaborating with **analyst network** Global Risk Insights.

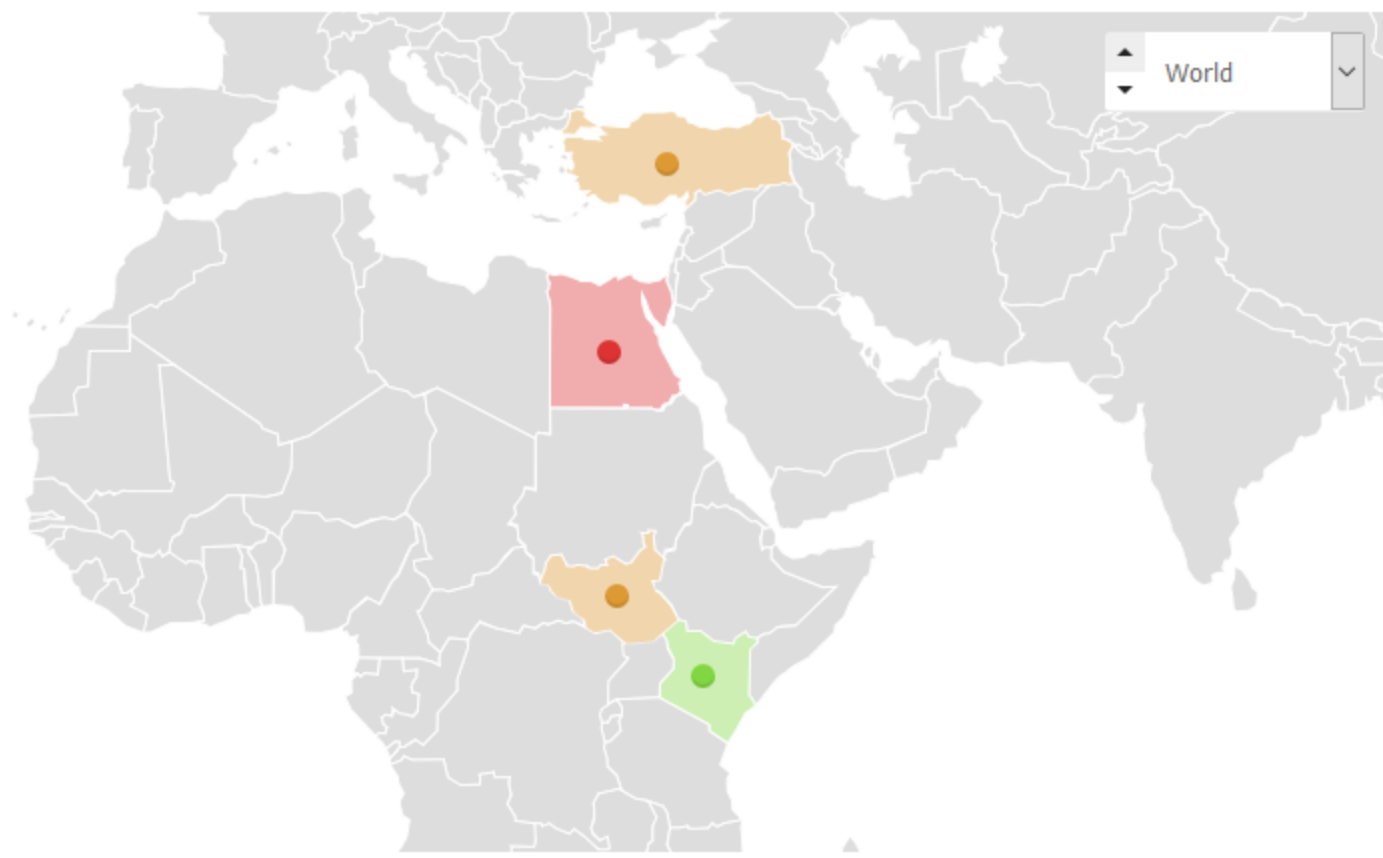
GRI has a team of 160+ analysts located around the world, including former presidential advisors, diplomats, and academics that:



GLOBAL RISK INSIGHTS  
Know Your World

- 1) Confirm/reject the predictions, thereby **increasing overall accuracy**.
- 2) Add a **further layer of analysis** to events likely to take place.

# Maps



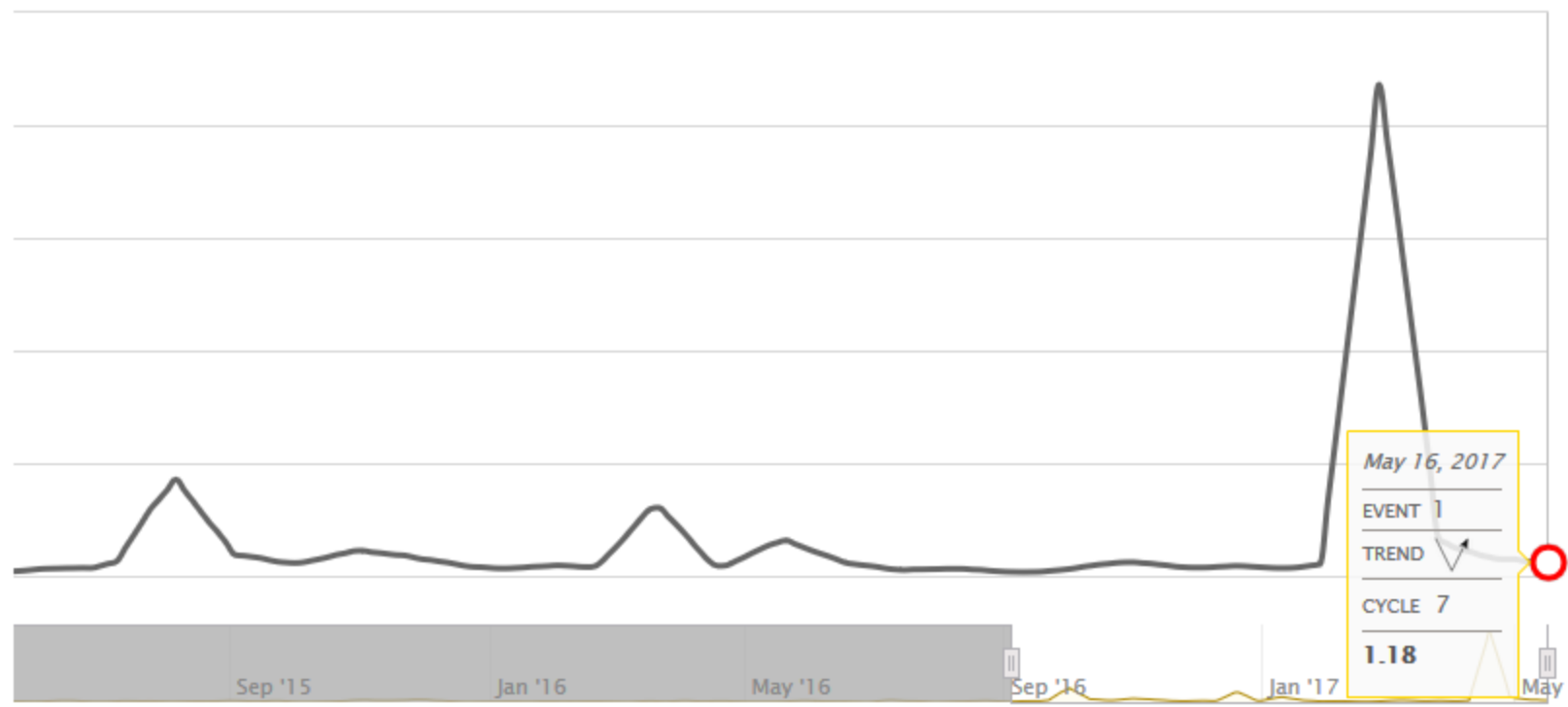
[Edit](#)

# Maps



[Edit](#)

# Egypt Media



May 16, 2017  
 EVENT 1  
 TREND ↘  
 CYCLE 7  
 1.18

LEGEND

1



# Egypt Media

## EVENT NARRATIVE

Sudan describes Egyptian presence in the disputed area of the Halayeb triangle as "occupation," wants International Court of Justice to oversee the issue. Over the dispute, the two countries have established trade restrictions, new immigration procedures, and have mobilized their militaries.

## STRATEGIC INTERESTS

Egypt opposes the creation of the Great Ethiopian Renaissance Dam on the Nile River. Egypt fears the possible impact of the construction of the dam on the Nile River's flow. Remembering that Egyptian politicians recklessly discussed sabotaging the dam in the past, Ethiopian and Sudanese authorities are on alert for a potential a military attack.

In addition, conflict in Sinai turning into a civil war, as clashes between local tribes and Islamic State are escalating.

## SOCIAL INFLUENCES

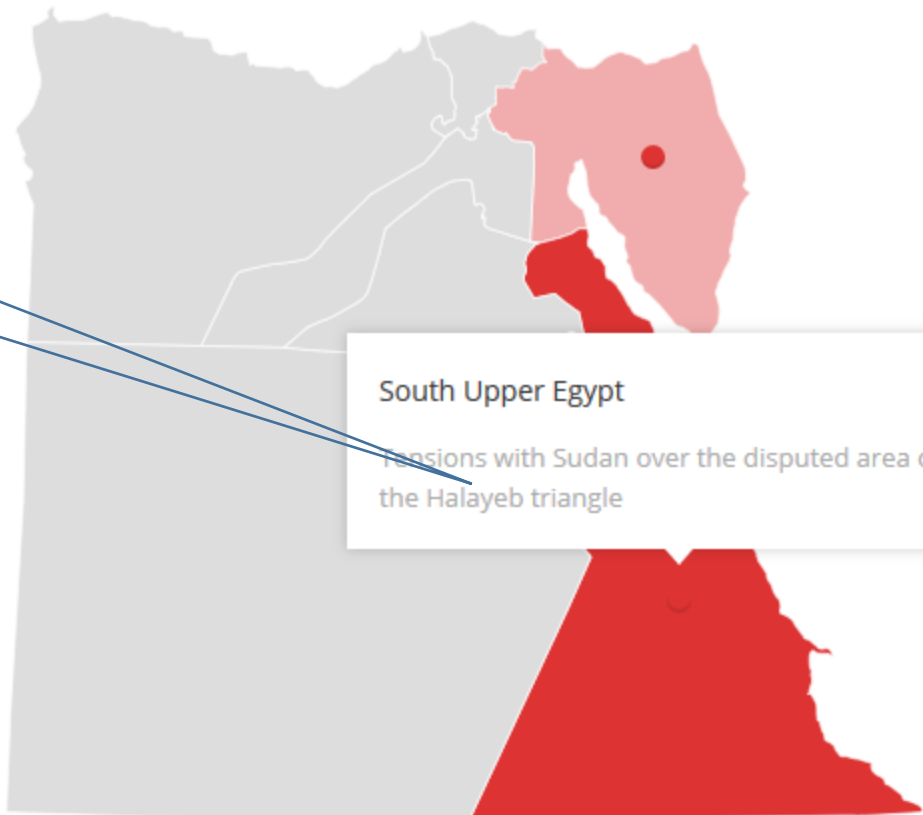
Sudan and Egypt have an agreement allowing citizens to move freely between both countries. However, Cairo has raised residency fees and denied entry to Sudanese citizens after the implementation of trade restrictions. It is estimated that about 850,000 Sudanese citizens reside in Egypt and many communities fear they will have to leave the country.

## ASSETS AT RISK

The recently increased animosity between the two countries contributed to the implementation of further trade restrictions on Egyptian agricultural goods by the Sudanese government. Sudan imports from Egypt about half a billion USD in goods, mostly food. Trade restrictions are causing concerns among the population, which relies heavily on the agricultural products imported from Egypt.

# Maps

Next, put the term "Halayeb Triangle" into the app, get a new prediction, accuracy increases.



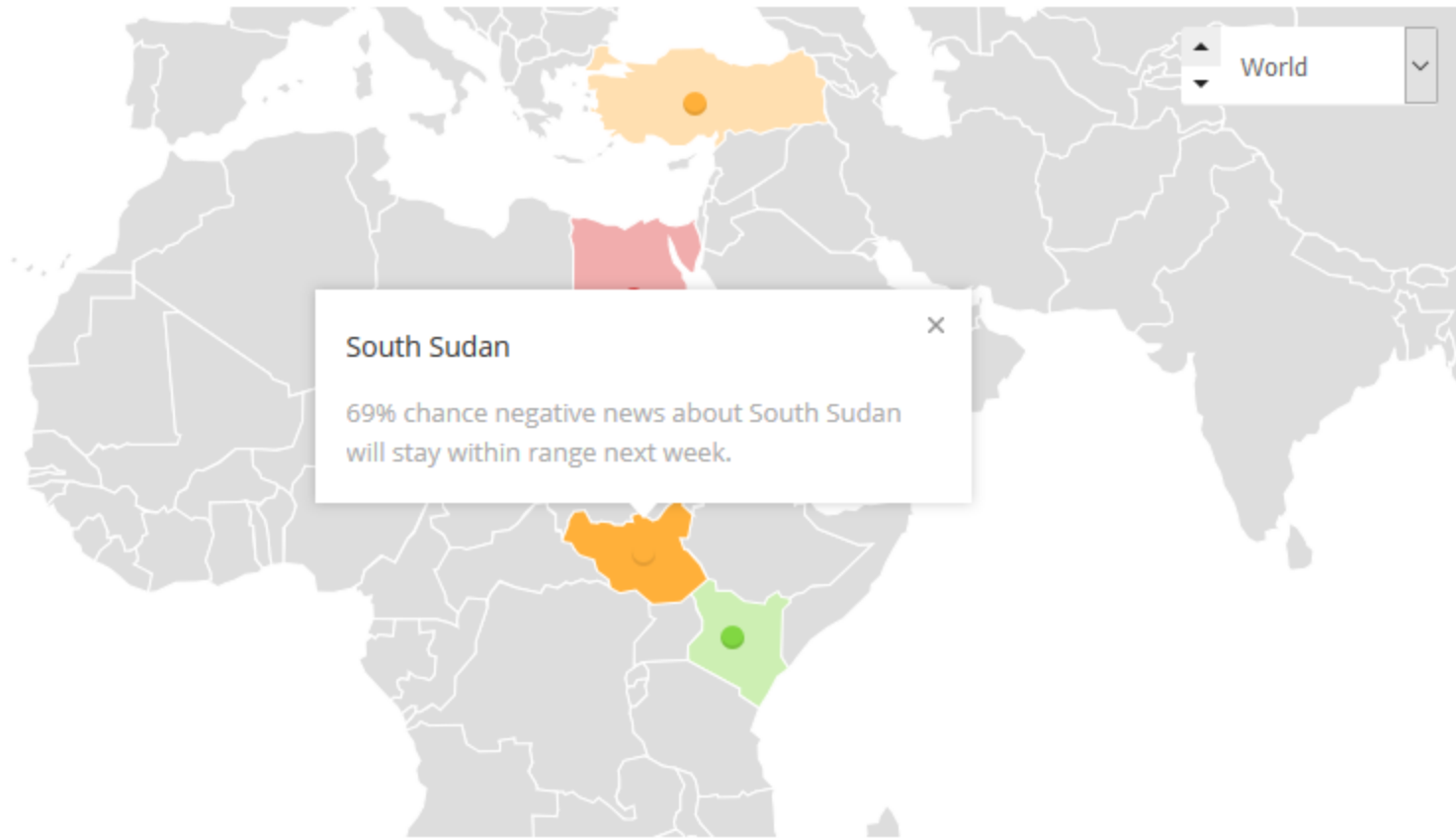
Egypt

South Upper Egypt  
Disensions with Sudan over the disputed area of the Halayeb triangle



Edit

# Maps



[Edit](#)





# 1



## South Sudan Media

### EVENT NARRATIVE

Combination of conflict, drought, and famine to continue state fragility.

### STRATEGIC INTERESTS

UNMISS deployed additional troops to Aburoc, in the Upper Nile region, to deliver assistance to over 50,000 displaced civilians and to provide assistance to civilians in imminent danger.

Ukraine accused of supplying weapons to South Sudan, despite ban.

### SOCIAL INFLUENCES

Cholera is spreading in Ayad County, in the Greater Upper Nile region. There is a growing concern regarding drinking water, as water-transmitted diseases, such as diarrhea could potentially kill thousands of people. Threat of famine increasing in the country.

South Sudan has been identified as "world's fastest-growing refugee crisis." Children most affected.

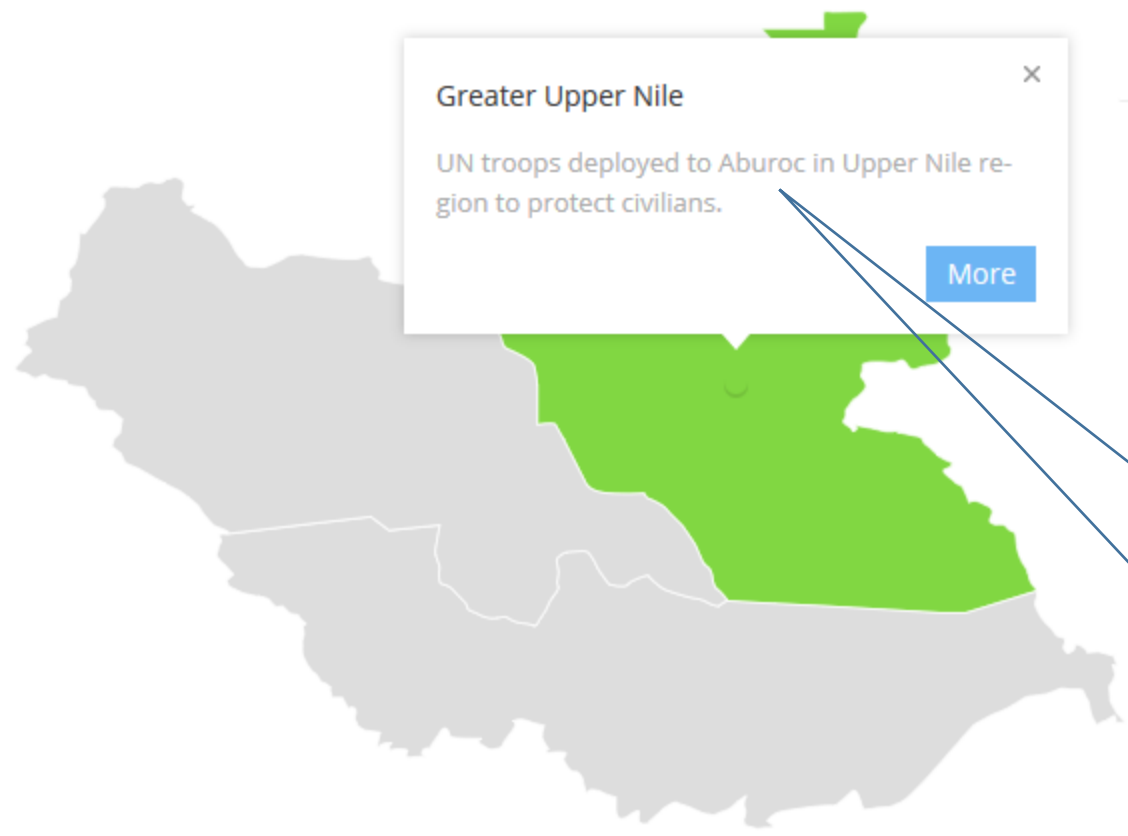
### ASSETS AT RISK

Shortage of funds to support humanitarian missions in the country is becoming critical.

Tweet Like 0 Share

Print This Page

# Maps



Greater Upper Nile

UN troops deployed to Aburoc in Upper Nile region to protect civilians.

More

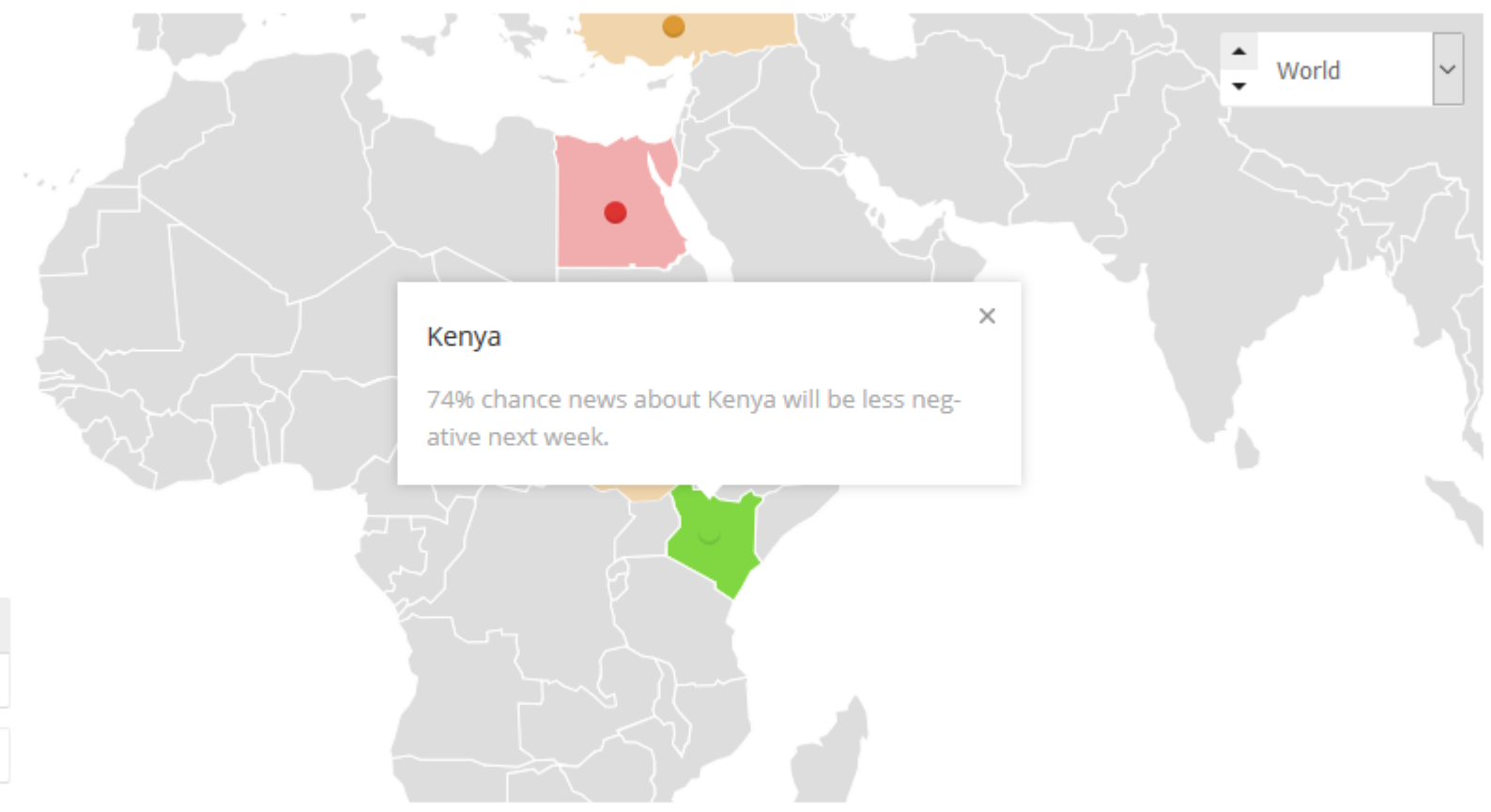
South Sudan



[Edit](#)

Next, put the term "Aburoc" into the app, get a new prediction, accuracy increases.

# Maps

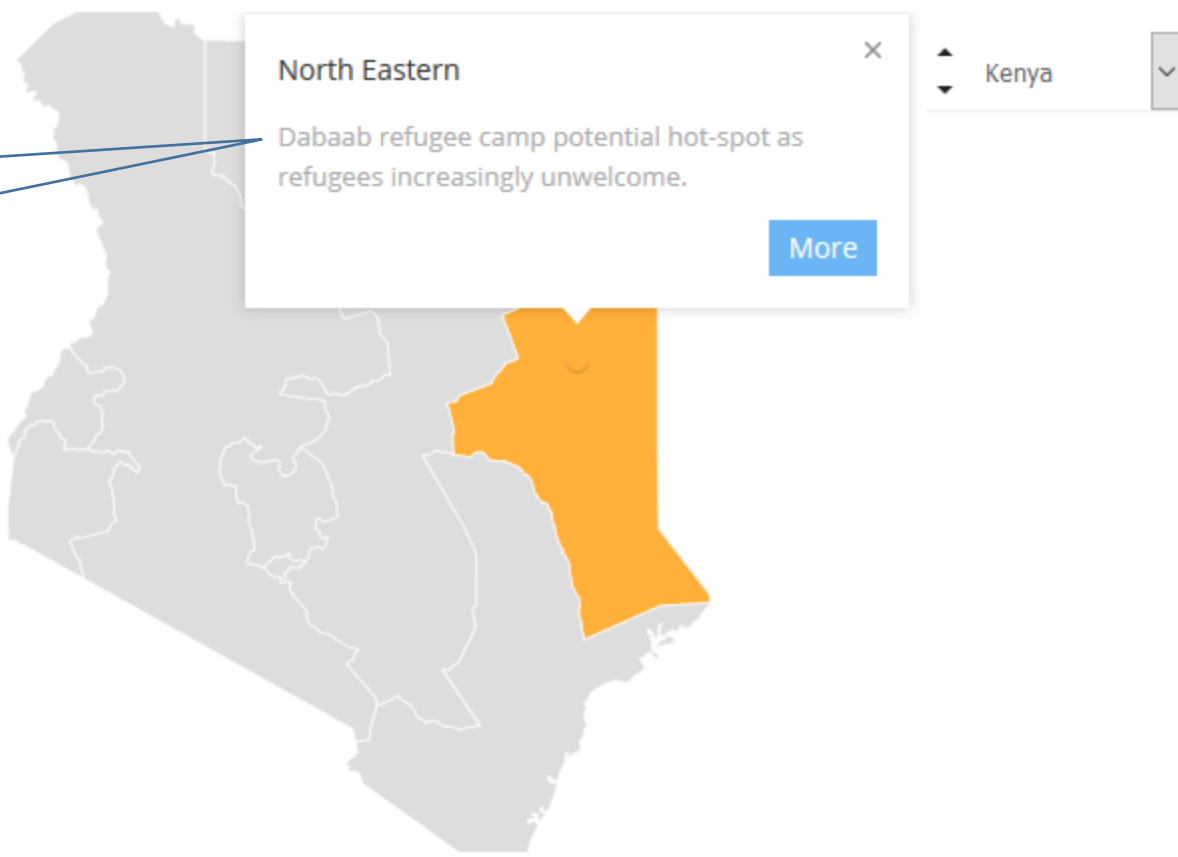


[Edit](#)





# Maps

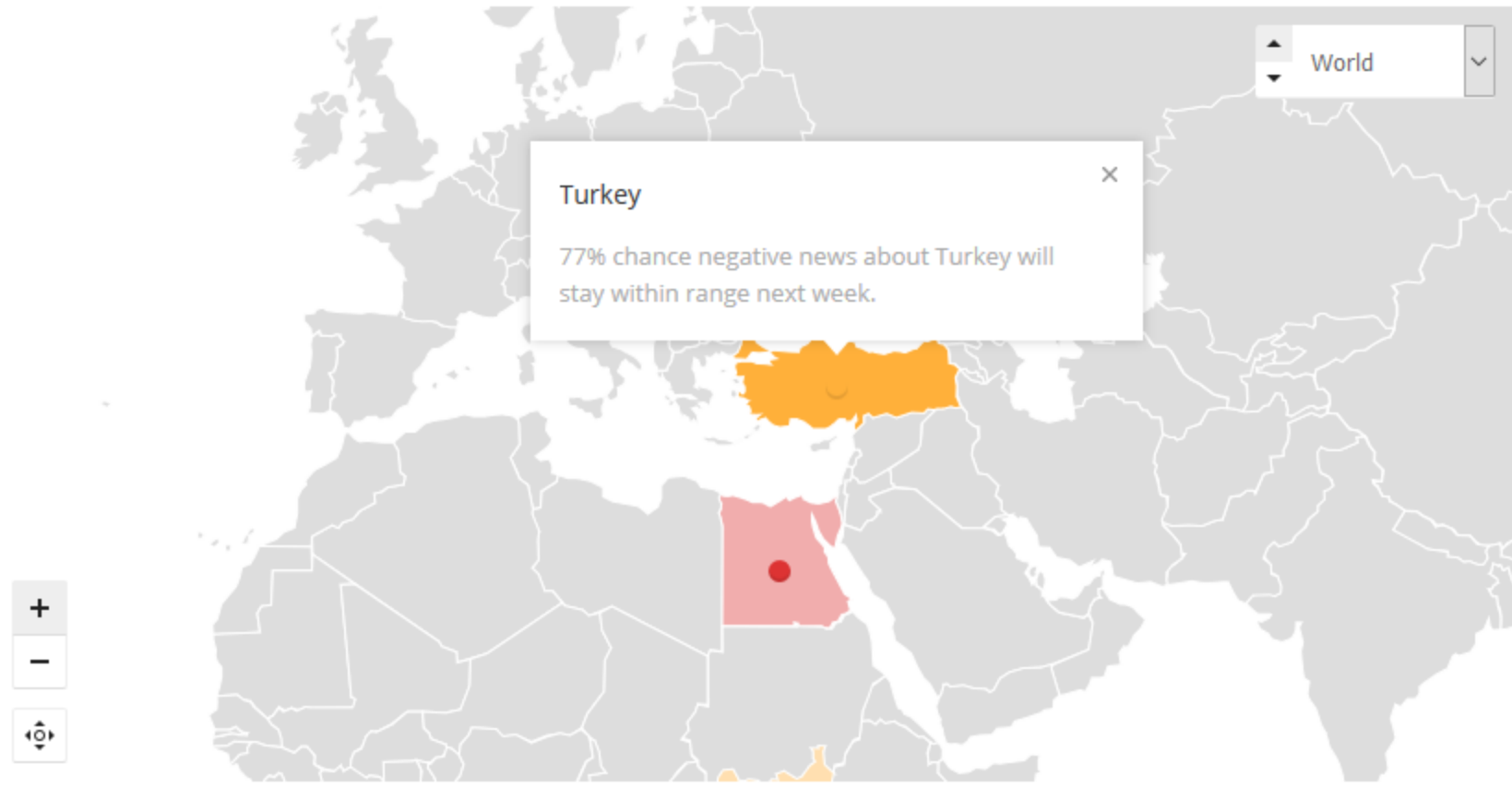


Next, put the term "Dabaab" into the app, get a new prediction, accuracy increases.



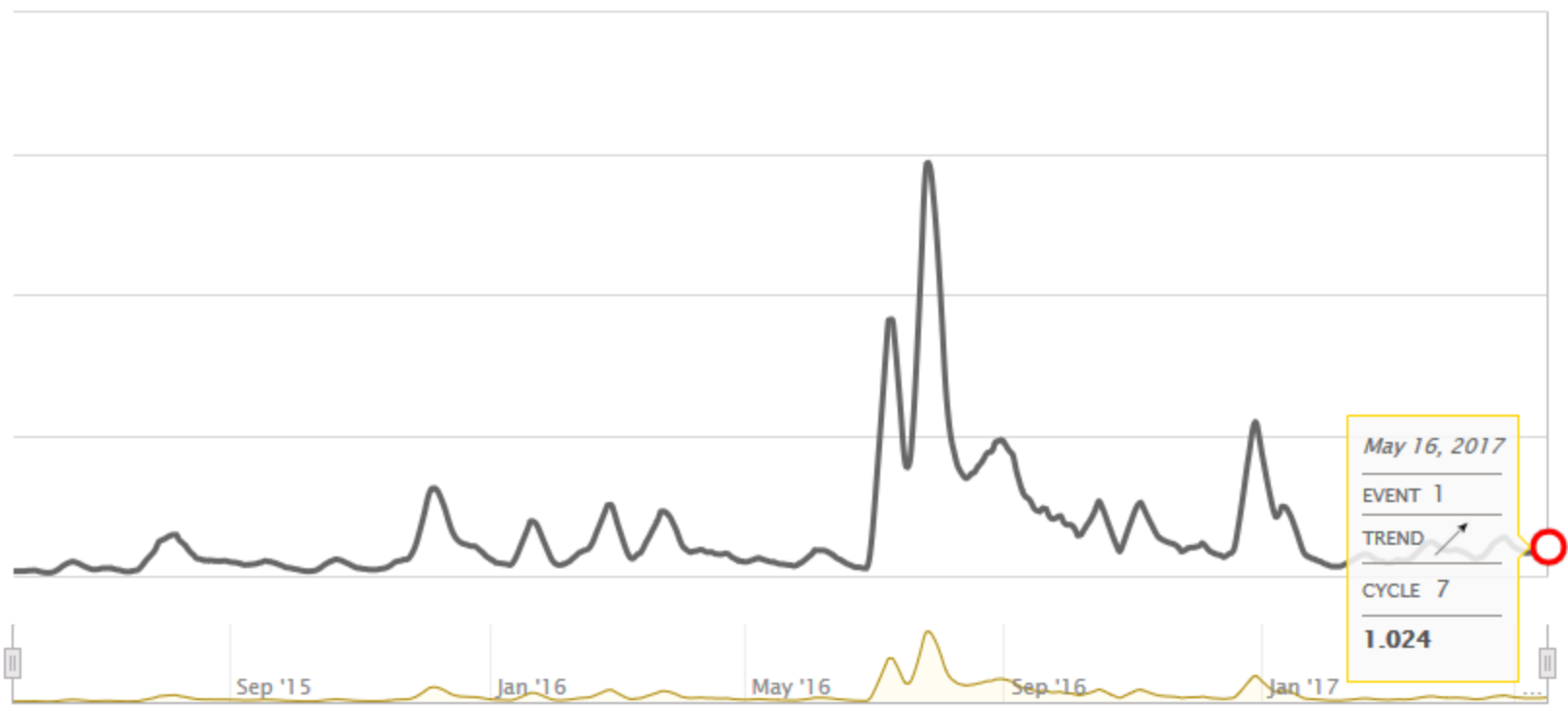
[Edit](#)

# Maps



[Edit](#)

# Turkey Media



LEGEND



# 1



## Turkey Media

### EVENT NARRATIVE

As the international community praises Turkey for taking three million refugees, allies worry about the deterioration of freedom of expression and other human rights, especially among the Kurdish population.

### STRATEGIC INTERESTS

Turkish-German relations continue to sour as Germany threatens to pull allied NATO troops out of Turkey.

### SOCIAL INFLUENCES

UN is investigating claims that the military killed hundreds of Kurds between July 2015 and December 2016 in the southeast.

### ASSETS AT RISK

As China keeps investing in Turkey, Turkey may become more assertive in disregarding UN regulations.

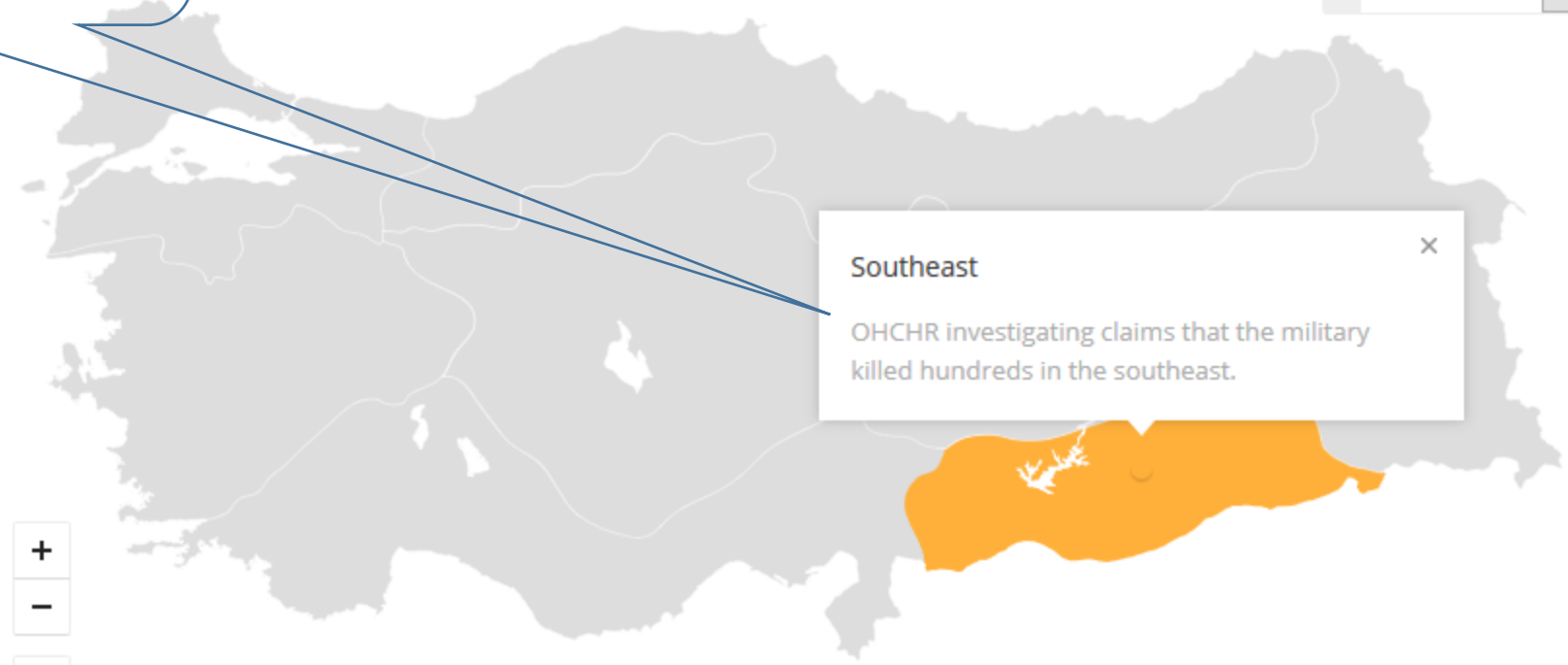
Tweet Like 0 Share

Print This Page

Next, put the term "OHCHR" into the app, geo-locate results from Turkey only, get a new prediction, accuracy increases.

# Maps

Turkey



**Southeast** [close icon]

OHCHR investigating claims that the military killed hundreds in the southeast.

## Conclusions from our research

To **increase accuracy** must include in the calibration of the model a detailed understanding of an organization's:

- Operations
- Priorities
- Goals\*

In sum: the more specific the inputs, the higher the accuracy.

---

Most accurate predictions concern future projections of a population's sentiment toward a specific project. This is achieved by **geolocation** of social media posts.