Internship Report (PeaceEye Project)

Opeyemi Kazeem-Jimoh

July-August 2023

spatial services

Co-funded by the Erasmus+ Programme of the European Union

COPERNICUS MASTER IN DIGITAL EARTH

PARIS LODRON UNIVERSITY SALZBURG

FFG Forschung wirkt.

Source: iStock

Introduction

Internship Objectives

(as outlined in my FemTech application)

- Literature review to determine state of the art in conflict management using Remote Sensing
- Literature review on existing conflict data and indicators. Also potential for integrating conflict data with other spatial datasets.
- User requirements gathering and mapping
- Current state of the art in satellite imagery delivery. (Subscription Platforms)

Workflow

1. Literature review
    - Methodology
        - EO based
        - Non EO based
    - Datasets
    - Data Integration Methods

2. Evaluation of Significant existing solutions (HIPE-2023, GeoQuery, ACLED Forecast Tool)

3. Potential Vendor Identification and Contact
    - Maxar SecureWatch
    - Airbus Living Library/One Atlas
    - Satellogic: Currently cheapest submeter imagery on the market.

Literature Review

Approach

WoS keyword search to identify 30 relevant Articles. In-depth analysis of 10 most relevant ones.

Keywords revolved around peace and conflict prevention with and without Earth observation data, conflict event forecasting methods e.t.c

Title	Author	Year
The Promise And Pitfalls Of Conflict Prediction: Evidence From Colombia And Indonesia	Bazzi.et.al	2022
While We Watched: Assessing The Impact Of The Satellite Sentinel Project	Raymond.et.al	2010
Remote Sensing For International Peace And Security: Its Role And Implications	Avtar.et.al	2021
Subnational Violent Conflict Forecasts For Sub-saharan Africa, 2015-65, Using Climate-sensitive Models	Witmer.et.al	2017
GEOQUERY: Integrating High Performance Computing (HPC) Systems And Public Web-based Geospatial Data Tools	Goodman.et.al	2018
Integrating Conflict Event Data	Donnay.et.al	2019
Utilizing Remote Sensing And Big Data To Quantify Conflict Intensity: The Arab Spring As A Case Study	Levin.et.al	2018
Comparing GDELT and ICEWS Event Data	Ward.at.al	2013
Using Machine-coded Event Data For The Micro-level Study Of Political Violence	Hammon and Weidmann	2014
Comparing Conflict Data	Raleigh.et.al	2019

Myth Busted: Contrary to claims by some sources, Climate change has not much to do with conflict occurrence or escalation: infact results prove the contrary.

This has been established in literature (Witmer et al, Raleigh et al).

Syrian conflict

Avtar et.al (2021)

Summary: Scoping review of role of RS/EO in Conflict, Peace and Security (major research concepts, and implementation of RS based techniques).

Methodology: Keyword search of extensive research databases using Latent Dirichlet allocation method

Datasets: Extant Literature

Key Findings
Not all operational aspects of armed conflict can be detected from EO-data, especially the factors that cause them. That being said, the following are an overview of the current role of RS and EO in conflict monitoring.

- **Refugee Monitoring and Relief (HR to VHR):** Movement detection, Camp expansion, Camp population estimation, Impact on environmental resources (Water, Land, Wood)

- **Armed group activities monitoring (MR - VHR):** Armaments (military instruments), rebel activity, atrocities monitoring, witness reports corroboration(burned area estimation using Optical imagery and MODIS thermal band).

- **Economic implication of armed conflicts (MR, VHR):** Changes in Nightlight pre and post conflict (VIIRS Black Marble), Building damage (including historical), Clandestine shipping activity (Piracy) monitoring through VIIRS Black Marble.

- **Conflict Prone Resource Monitoring:** Rebel activities around resources (also related to rebel financing), Monitoring resource based disputes to prevent full-on armed conflict, resource exploitation activity around contested regions.

- **Challenges to Legal Application:** It is important to note that the challenge of universality and standardization, in the interpretation of EO data, is a major factor in its global acceptance particularly in the eyes of the Law.

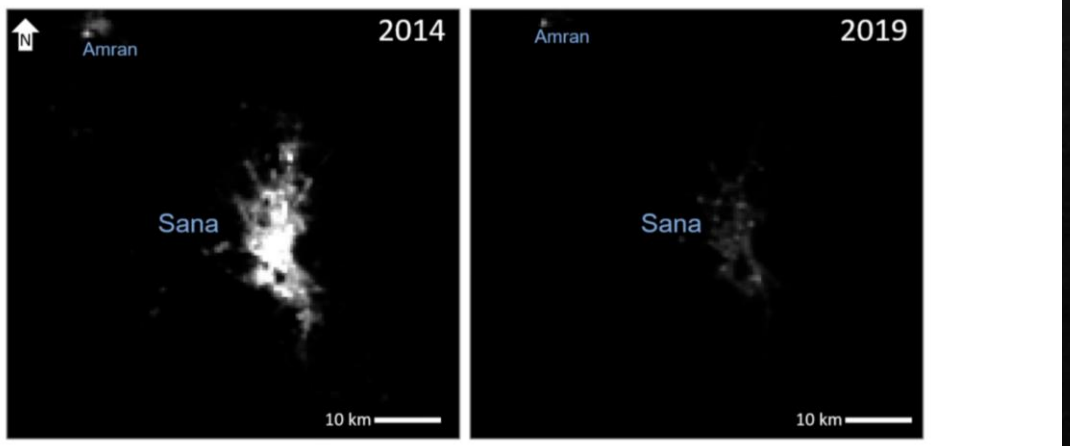


Figure 6. The Suomi National Polar-Orbiting Partnership Visible Infrared Imaging Radiometer Suite sensor (NPP-VIIRS) night-time light images of Sana, Yemen, during the years 2014 and 2019.

Bazzi.et.al (2022)

Summary: Using unusually fine-grained historic conflict, geographic and socio-economic data, they attempt to predict violence one year ahead with a range of machine learning techniques.

Methodology: Machine learning (Ensemble Model made up of 4 different techniques (Random Forrest, Gradient Boosted Machine, LASSO, Neural Networks)).

Datasets: High resolution historic data as the dependent variable (SNPK, CERAC, Podes), 482 subdistrict-level predictors/covariates (e.g. population, religion, remoteness, sector share, demographic, geography, economic output, natural disasters etc.) split into time variable and time invariable factors.

Key Findings:

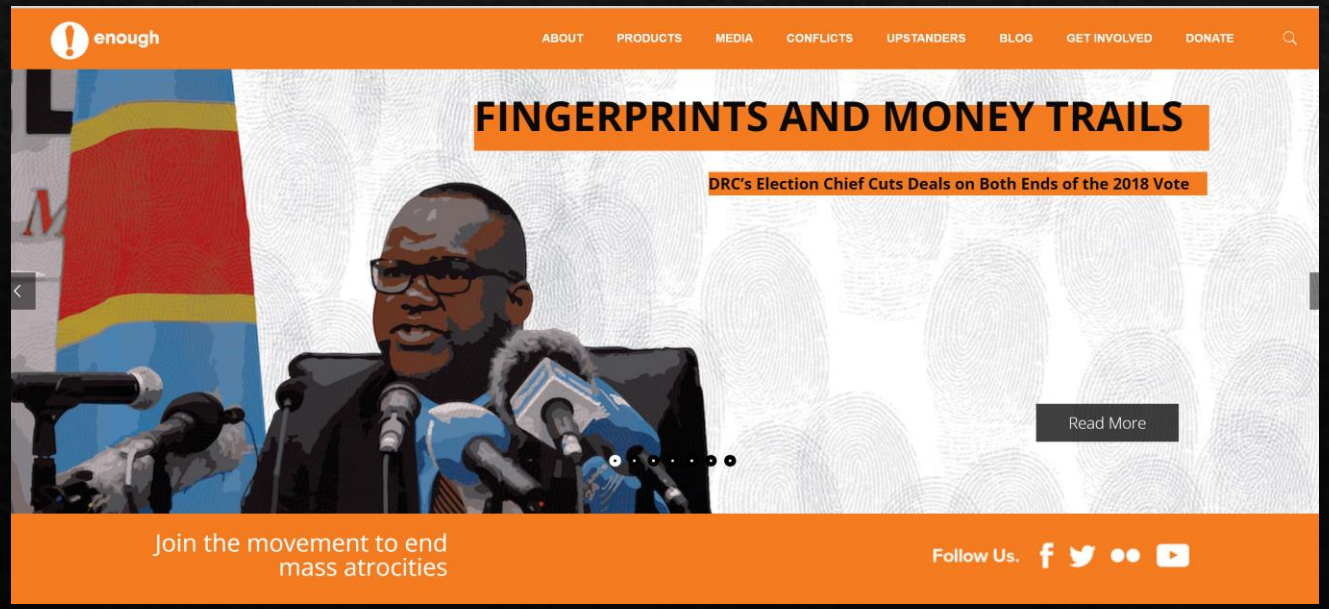
- Lagged dependent variable alone(whether or not conflict offered in the area in the past), is not effective for prediction. Other important details like severity of particular incidents, identification of actors involved etc. prove effective.
- Model performed better in Conflict prediction over space than time.
- Time invariable factors (terrain ruggedness, share of local economy, distance from capital, etc..) have more predictive capacity than time-varying (fast-onset disasters, economic output, government policies, etc.) factors.
- Extensive covariates alone predict well.

Raymond et.al (2010)

Summary: Through consistent border monitoring, using VHR EO data and analytics, identify indicators for mass atrocities and influence public policy to prevent further escalation of conflict.

Methodology: Citing the need for consistent accurate interpretation of imagery, in-house standards and thresholds for certainty were developed for the identification and classification of indicators of mass-atrocities.

Datasets: Near-realtime VHR imagery over the AOI (border) during the pilot phase (11 months) provided by Digital Globe (Maxar).



Key Findings:

- With better standards developed, approach shifted from mere documentation to prediction, enabling more predictive image tasking.
- Evidence based proof (usually made possible by EO data), of concerning activities brewing is not enough to prevent escalation (as echoed by Witmer et.al 2015).
- Even though the analysis carried out by the HHI, the inclusion of high-profile individuals in the organization may have damaged the credibility of the unbiased results being produced.
- Their developed methods are not publicly published.

Witmer et.al (2017)

Summary: Extensive statistical modelling of past and current climate(global rainfall and temperature models), conflict data (ACLED) and covariates (population, IMR, governance(regime type), distance to border/capital city) to form a baseline model.
Then using future projected values for climate and covariate information, predict the future conflict situation, particularly withing the context of Shared Economic Pathways (SSPs)

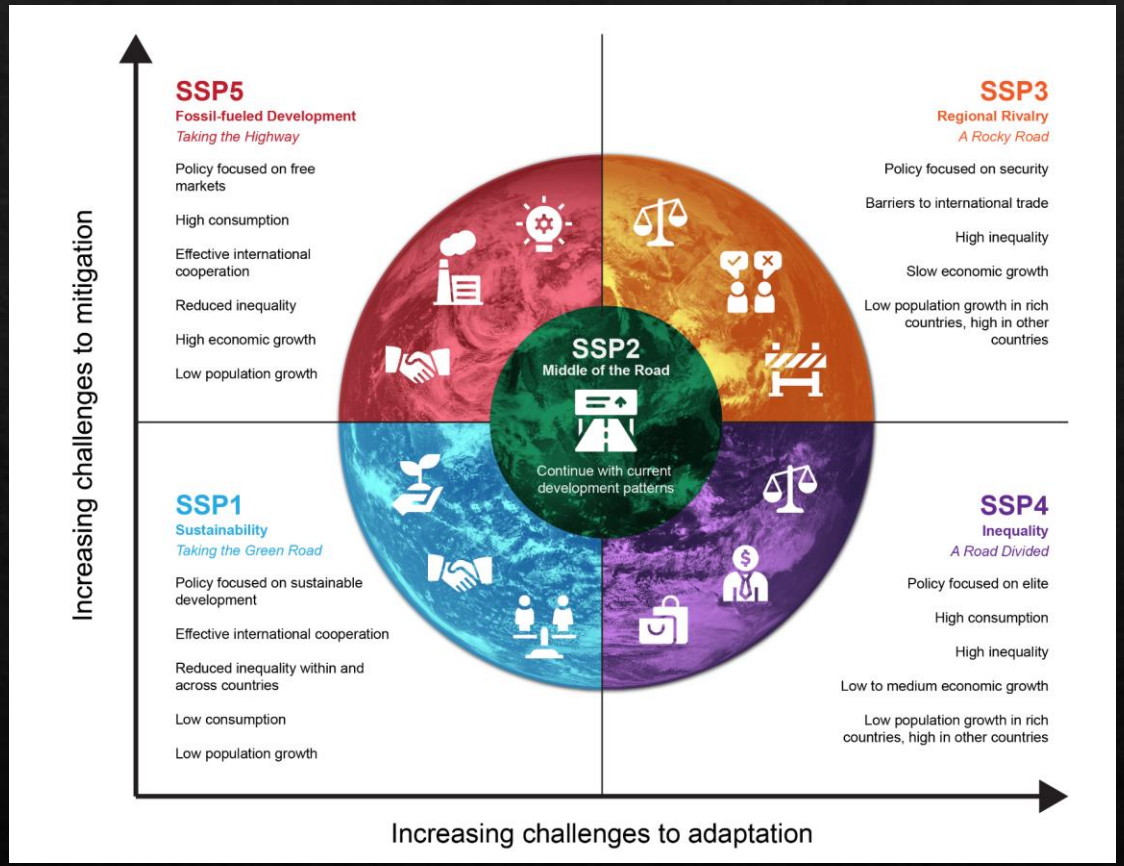
Methodology: Statistical analysis (AUC-ROC, AUC-PC)

Key Findings:

- While there is statistically significant relationship between amount of violent conflict, political rights, population size and rising temperatures…
- Future violence is very sensitive to changes in political rights and largely insensitive to temperature or rainfall anomalies.

- For forecasts of violence under pessimistic future scenarios of unchanging political rights from current levels (SSP3 and SSP4), rising temperatures and increasing population exacerbate current levels of violence…
- However, If political rights and governance do improve (SSP1 and SSP5), then future levels of conflict are likely to remain stable or even decline inspite of increasing temperature and population.

- **Views echoed by** Clionadh Raleigh (ACLED CEO).
- **"Conflict is just a power play in developing conflict-prone regions"**



Shared Socioeconomic Pathways

Raleigh et.al (2019)

Summary: A comparison of GDELT, UCDP GED, ICEWS, GTD, Phoenix event dataset and ACLED in terms of the inclusion criteria, methodology, and sourcing.

Human-coded (HC) datasets:
- Armed Conflict Location & Event Data Project (ACLED)
- Uppsala Conflict Data Programme Georeferenced Event Dataset (UCDP-GED)
- Global Terrorism Database (GTD)

Machine-coded (MC) datasets:
- Global Database of Events, Language, and Tone(GDELT)
- Integrated Crisis Early Warning System(ICEWS)
- Phoenix event dataset

Key Findings:
- HC datasets are ultimately superior to MC in terms of data quality.

- Though they may be updated more frequently and feature some advance ML techniques under the hood, MC datasets are found to contain irrelevant and misleading information, event duplication and redundancy.

- ACLED performs better than other HC due to:
  - Less exclusion criteria for armed conflict events (no minimum fatality, no strict definition of conflict event types)
  - Rigorous and locally tailored source and review approach, transparent source inclusion criteria (back tracking to source).
  - Spatial and temporal precision transparency.

Donnay et.al (2019)

Introducing MELTT

An open-source tool for seamlessly integrating different conflict event datasets.

Challenge addressed: Same events may be captured differently in distinct datasets, raising the issue of duplication and redundancy. This tool provides a systematic approach to evaluating multiple occurring events, and aggregating them into one entry in the final combined database.

Methodology: Leveraging the assumption that events not occurring in the same time and space (barring some fuzziness) cannot be the same, a 4-step process is employed which include;

- Spatio-temporal Blocking (semi-automatic)
- In-depth comparison (manual)
- Isolate Matching entries (deferred-acceptance algorithm)
- Merging matching pairs into a single entry.

Testing done in Nigeria 2011, South Sudan 2015, Libya 2011.

Notes

- The integration protocol is sensitive to data quality, alongside the spatio-temporal windows specified by the researcher.

- The framework also assumes that the researcher has already preprocessed the data to ensure that entries across data sets are indeed comparable in these respects (taxonomy)

Key findings

- The degree of precision (either exact geocode of event locale, or just a point in the center of the admin boundary) in data reporting is indicative of the quality of reporting, with more precise geocoding noticed in datasets with wider coverage and completeness.

Package 'meltt'

October 27, 2022

Type Package

Title Matching Event Data by Location, Time and Type

Version 0.4.3

Date 2022-10-22

Author Karsten Donnay and Eric Dunford

Maintainer Karsten Donnay <kdonnay@gmx.net>

Description Framework for merging and disambiguating event data based on spatiotemporal co-occurrence and secondary event characteristics. It can account for intrinsic ``fuzziness'' in the coding of events, varying event taxonomies and different geo-precision codes.

License LGPL-3

Depends R (>= 2.6)

Imports utils, ggplot2, plyr, dplyr, reticulate, Rcpp, shiny, shinyjs, tidyr, leaflet, tibble, methods

SystemRequirements Python (>= 3.6)

NeedsCompilation yes

LinkingTo Rcpp, RcppArmadillo

Repository CRAN

Date/Publication 2022-10-26 22:22:36 UTC

R topics documented:

meltt-package . . . 2
crash_data1 . . . 3
crash_data2 . . . 3
crash_data3 . . . 4
crash_taxonomies . . . 5
is.meltt . . . 5
meltt . . . 6

Matching Event Data by Location (meltt) R software package that implements the full protocol is freely available from CRAN (See documentation).

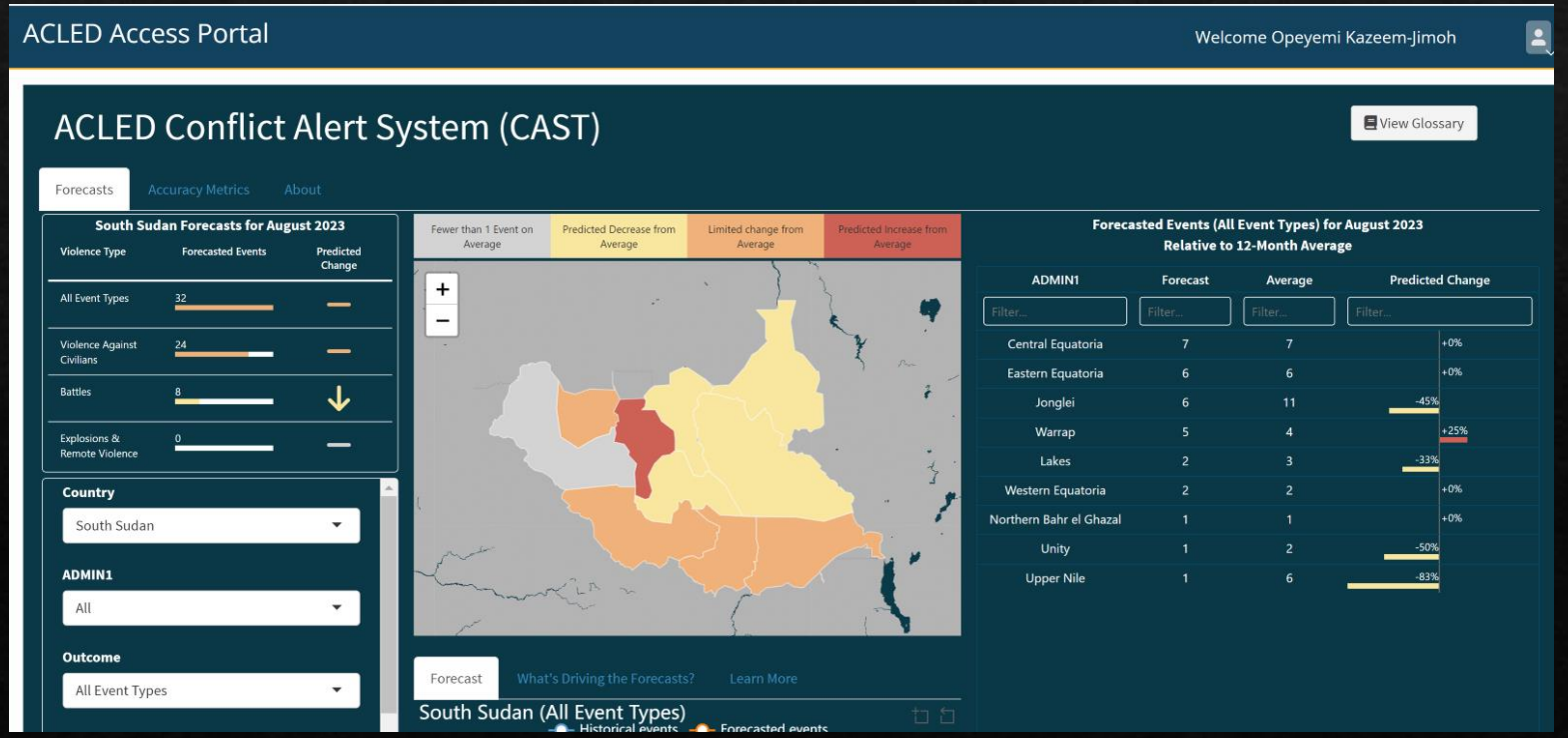
Literature Review Document

More Existing Solutions

GeoQuery (Next slide)

DRC-Forecast Tool, Balcony (HIPE2023)

ACLED Early Warning Systems – Early Warning Dashboard, Conflict Alert System (ACLED Primer 7)



GeoQuery

An Online Web Platform through which users request integrated geospatial data spanning geophysical, environmental and socio-economic measurements, aggregated to user specified administrative boundaries (up to subnational).

Results are in form of permanent links through which users can access their data (in .csv format) and accompanying metadata.

Methodology:

- **Major challenge overcome:** Each data source require customized sometimes extensive processing pipelines, before datasets are ingested into GeoQuery.

- Provides of an extensive list of currently relevant conflict, socio-economic and demographic datasets.

What Data is Available Through GeoQuery?

Below is an overview of available data. For more information, visit AidData's complete catalogue of data available through GeoQuery.

Boundary Data

GeoBoundaries (William and Mary), Global Grid 0.5 decimal degrees, Chinese financed development project sites, major cities, and more

Measurement Data

International Aid

Multiple data sources from the AidData research lab, including World Bank and Chinese development projects, country-specific (e.g. Afghanistan, Nepal) datasets.

Population and the Environment

Population Density and Counts (CIESIN), Slope and Elevation (NASA), Protected Areas (IUCN), NDVI (UMD GLCF), Land Cover (European Space Agency, NASA), Precipitation and Temperature (UDEL)

Conflict and Health

Conflict deaths (UCDP), Conflict Events (ACLED), Lootable Gold Deposits (GOLDATA), Child Mortality (Stanford), Ozone Concentration and Particulate Matter (TM5-FASST).

Economic Development

Nighttime Lights (DMSP; VIIRS), On-shore petroleum (PRIO), Gemstone Deposits (GEMDATA), Gross Domestic Product (CIRES), Drug Cultivation Sites (DRUGDATA)

Access to Infrastructure

Distance to Coastal features and Water (GSHHG), Distances to Roads (gRoads, CIESIN), Distance to country borders (GADM), Travel Time to Major Cities (JRC)

Datasets & Indicators (OS-INT)

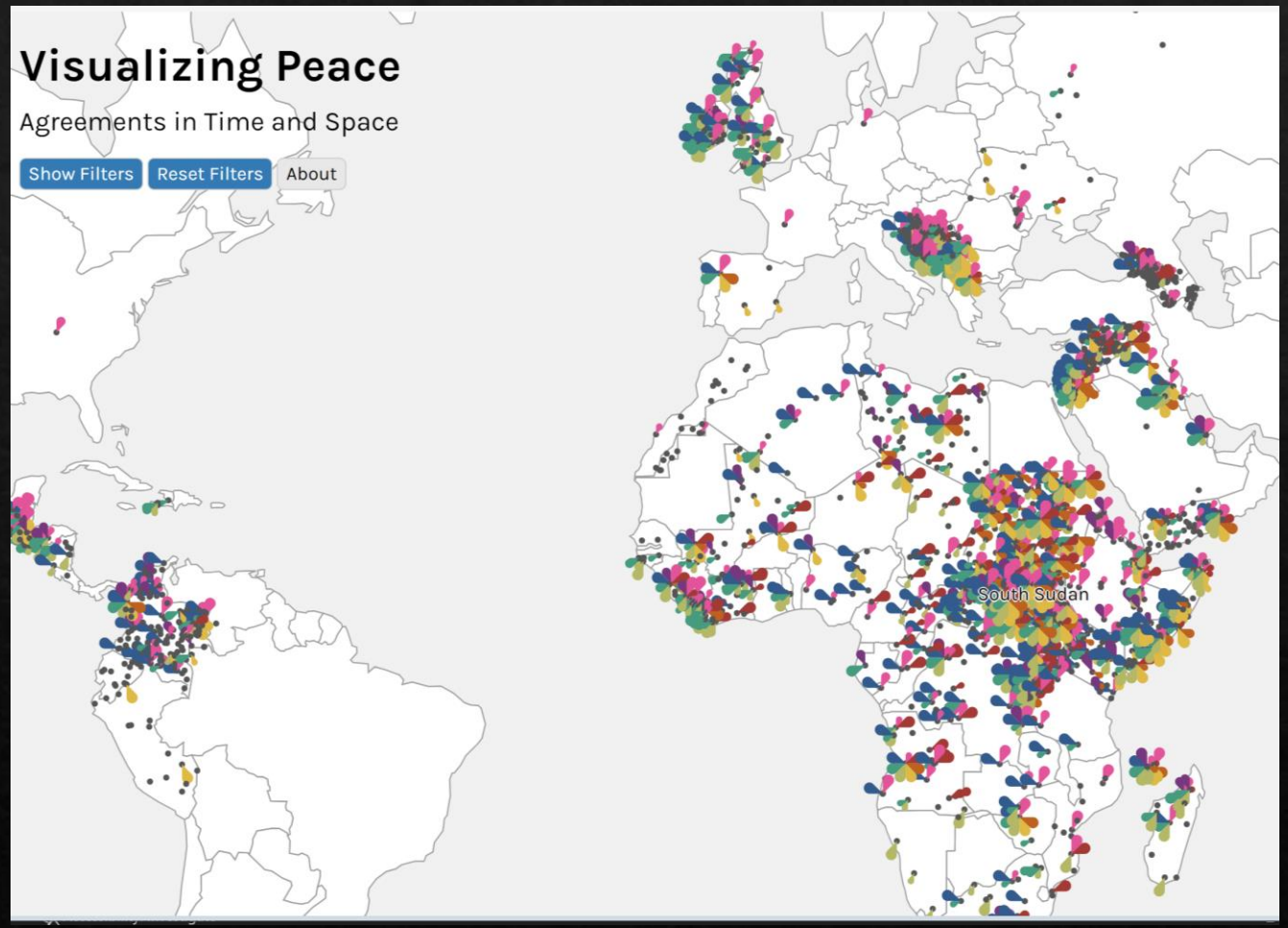
Conflict Data:

- ACLED (ACLED primer, Raleigh et.al, Donnay et.al)
- Others: GDELT, ICEWS, Phoenix, UCPD-GED, GTD (Ward.et.al 2013, Hammon and Weidmann 2014).

Socioeconomic Indicators and Dataset:

- Food insecurity: DIEM (DIEM report)
- Historical Election Data (Global Election Database, Constituency Level Election Data (CLEA))
- Election Violence Vulnerability Index (Koffi Annan Foundation)
- Peace Treaties: Peace Agreements Database (time and space) by University of Edinburgh
- **Famine Early Warning Systems Network(FEWS – South Sudan)
- Infant Mortality Rate (UN, World bank)
- Future Election Data (None; but can be created (next slide)---

E.g. South Sudan Elections

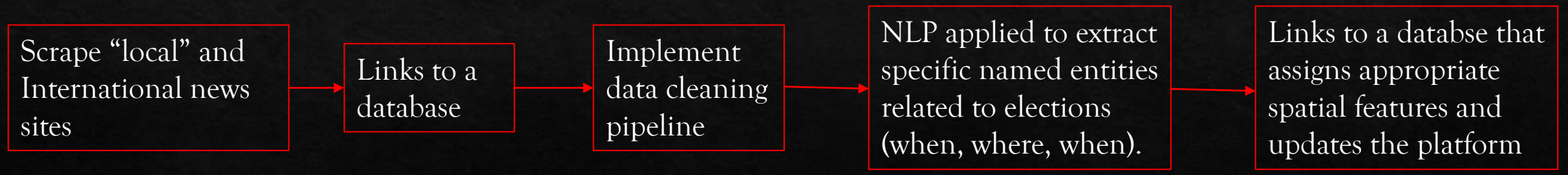


Hands On (NLP/Webscraping)

With the aim of creating a Web enabled Database on Potential Elections data (Case study South Sudan)

News Sources (Al-Jezeera, CNN, BBC)

Limitation
Not as straight forward as initially thought. Such datasets may be assessable from other organizations currently undertaking it.

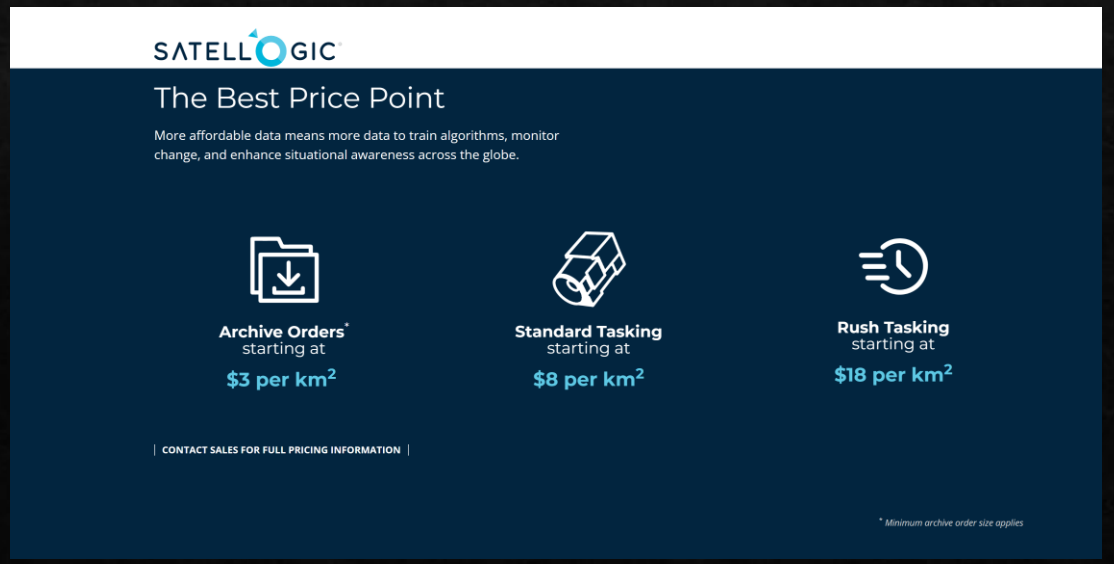


Proposed Workflow

Potential Vendor Identification

Maxar Secure-Watch (Secure Watch Primer, Table)
Airbus Living Library (Price Sheet, Table)
Satellogic (Website)

		Maxar Secure Watch	Airbus One Atlas(Living Library)	Planet	Satellogic (Aleph) *beta
	Functionality	Online and Archive (Worldview 1 to 3, GeoEye-1), Archive only: (Quickbird, WV 4)	16 Optical (Including SPOT, Pleiades, Pleiades neo, Vision-1, DMC) 3 Radar (TerraSar-X, TanDEM-X, PAZ. Radar not included in Living Library) Living Library contains the Archive iata and it is what is available for subscription. OneAtlas	PlanetScope 3.7m (0-5m) Dove Constellation(C, R, Superdove), SkySat (50cm for tasking only), RapidEye(5m). (No response recieved from Sales team for enquires)	Aleph-1 (40+ OneSat) Satellite Constillation, 300 planned. 0.7m res MS imagery and 25m multispectral (They do not have a publicly available subscription platform, but they are included here because they have currently the cheapest
1	Most recent imagery visible	Yes (also as base layer)	Yes		
2	Quality of most recent imagery	best of the best (max 15% cloud cover)	best of the best (at most 15% CC (Neo), 25% CC (Pleiades 1&2). 0% SPOT, 30% incidence angle (Pleiades), 25% SPOT, Orthorectified)		
3	Tiered Subscirption/Pricing	yes (Standard, Premium), GB subscription (the more GB subscribed, the lower teh cost).	no		
4	Subscription Validity	yearly	yearly		
5	Top-up	possible (not automatic)	possible (not automatic)		
6	Minimum subscription	5,000 euro	5,000 euro		
7	quota tracking	available	available		



Conclusion

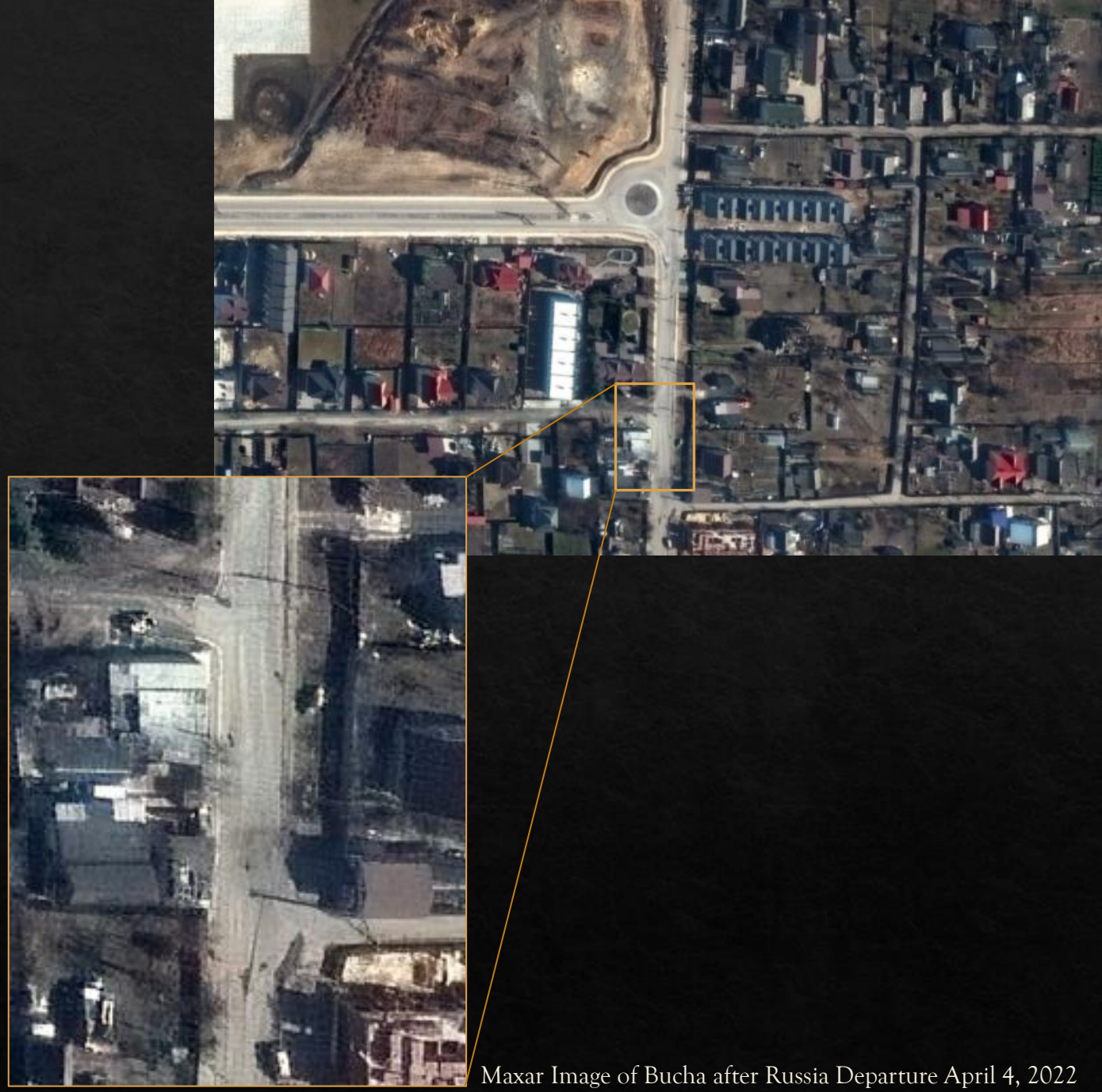
Conflict response is shifting from monitoring and evaluation to forecasting and early warning systems, a line which is and should be followed by PeaceEye.

The methods in current state-of-the-art in conflict forecasting, while still experimental, involves using disaggregated conflict and other socioeconomic data and advanced statistical methods and ML. However, RS and EO still has a role to play.

Combined with HumINT (local peace activists), other socioeconomic data not currently captured (e.g. potential election data including dates, renumeration areas…etc.), PeaceEye could have the upper edge by improving on the currently existing solutions.

Appreciation
Going Forward **(User requirements Work Package)**
Q&A…

Vielen Dank!!!



Maxar Image of Bucha after Russia Departure April 4, 2022