Enhancement of Project AEDES

Inception Report for

UNICEF Philippines and UNICEF Digital Public Goods Alliance

Prepared By  
Cirrolytix Research Services  
January 2022



Applied Analytics for Social Impact

THIS PAGE INTENTIONALLY LEFT BLANK

# 

Table of Contents

[Introduction 6](#_Toc93337139)

[Overview 6](#_Toc93337140)

[About Project AEDES 6](#_Toc93337141)

[Assessment of Current State 6](#_Toc93337142)

[Objectives of the Project 6](#_Toc93337143)

[Enhanced AEDES Platform 8](#_Toc93337144)

[Technology Stack 8](#_Toc93337145)

[Data Architecture 8](#_Toc93337148)

[Data Sources 8](#_Toc93337149)

[Licensing 9](#_Toc93337150)

[Data Model 10](#_Toc93337151)

[High-Level Data Architecture 10](#_Toc93337153)

[Datasets 10](#_Toc93337155)

[Data Collection 12](#_Toc93337156)

[Risk-based Framework 13](#_Toc93337157)

[Use-Cases 13](#_Toc93337160)

[INFORM AEDES Data 13](#_Toc93337161)

[Project Plan 14](#_Toc93337162)

[Expected Outcomes 14](#_Toc93337163)

[Risks 14](#_Toc93337164)

[Project Activities 14](#_Toc93337165)

[Project Design 16](#_Toc93337166)

[Organizational Profile 18](#_Toc93337167)

# 

Introduction

Overview

In 2019, the Philippine government declared a national epidemic as it struggled to contain its worst dengue outbreak since 2012. Total cases reached 420,453 and deaths at 1,565, 78% increase in infections and 33% in deaths from 2018. 42% of deaths are children between 5 and 9 years old (Relief Web, 2019) and 2 more kids die of dengue fever in Zambales (Inquirer, 2021).

With manual reporting, release of data is delayed which hampers the health sector’s ability to effectively deal with the threat. Prioritizing prevention is essential.

There is a risk that the dengue epidemic continues to take lives undetected. According to a study by Seposo, a decrease could be from reporting hesitancy due to fear of contracting COVID-19 in a health facility. COVID-19 may recede after mass vaccinations; however, dengue will continue to kill.

About Project AEDES

Project AEDES was developed in 2019 as a big data early warning and surveillance system for dengue. The system intends to nowcast probable dengue cases and dengue-related deaths using Google Search Trends, precipitation, and temperature readings from climate data. Google Search Trends represent the public interest and panics related to dengue. Moreover, to detect potential mosquito hotspots, the system utilizes FAPAR, NDVI, and NDWI readings from remote sensing. Information derived from such data would help prioritize interventions and resource allocation.

Assessment of Current State

Since the platform was developed under the NASA International Space Apps Challenge, there are several known limitations and opportunity for enhancements, which includes the following:

* Manual data gathering and preparation
* Unavailability of current data
* Counter-intuitive dashboard for further elaboration
* Unclear policy usage of the information
* Lacks relevance due to limited scope of solution (i.e. dengue)

Objectives of the Project

Following the limitations stated above, the key improvements to the AEDES functionality are:

* Automation of data gathering from various sources, especially weather data which was sourced from an offline source (DOST-PAGASA)
* Addition of new weather, satellite, geospatial and socioeconomic data to enrich dataset
* Enhancing the predictive modeling by adding additional ML algorithms to improve model fitting performance
* Incorporating the INFORM Epidemic Risk Framework with data gathered by AEDES teams to generate location-based risk maps, and advise policy interventions to mitigate the impacts of dengue
* Improvement of User Interface to make it feel more like a consumer utility (e.g. Waze)

Enhanced AEDES Platform

Technology Stack

Graphical user interface, application, logo, company name

Description automatically generated

Data Architecture

Data Sources

Licensing

Open Data

* Satellite Data - Landsat 8 (NASA) and Sentinel-2 Copernicus (ESA)
* Wunderground Weather Data - Wunderground
* Google Search Trends - Google https://policies.google.com/terms
* Twitter Data - Twitter
* OpenStreetmap Data - OpenStreetmap
* Disease Surveillance Data - Department of Health (via Freedom of Information)

Freemium

* Accuweather Weather data - Accuweather

Data Model

Diagram

Description automatically generated

High-Level Data Architecture

Datasets

Project AEDES relies on the following datasets as input data to the dashboard:

* Google Trends Data

The raw values, ranging between 0-100, represent search interest as the proportion of all searches of the inputted term for a given period of time and location. A value of 100 is the peak popularity of the term, 50 represents half of the popularity, and 0 means inadequate data.

Search data of related dengue words from Google Trends are collected at specific areas of interest using dengue top 5 dengue keywords as dictated by Google search trends.

* Disease Surveillance Data

Monthly reported dengue cases and deaths of the selected regions from 2015 to 2018 were obtained from the public records of disease surveillance released by the Department of Health on their website.

* Weather Data

Time-series data of multiple weather attributes such as temperature, precipitation rate, accumulation of precipitation, wind, solar radiation, etc are collected from Wunderground and Accuweather.

* Satellite Data

Process Sentinel2, Landsat, and MODIS band satellite data using python to determine stagnant water locations (NDW), high vegetation (NDVI), high built-up index (NDBI), air quality index (AEROSOL) and solar irradiance/surface temperature in the specified area of interest. The generated dataframe coinciding specific pairs of longitude and latitude are used to determine hotspots of dengue outbreak.

* Twitter Data

Scrape dengue-related twitter data in order to capture time-series tweet count, engagements such as retweets and likes, and tweet sentiment such as positive, negative, or neutral.

* OpenStreetMap Data

Availability of clinics and hospitals in an area are available through OpenStreetMap. Data include distance and count from points of interest.

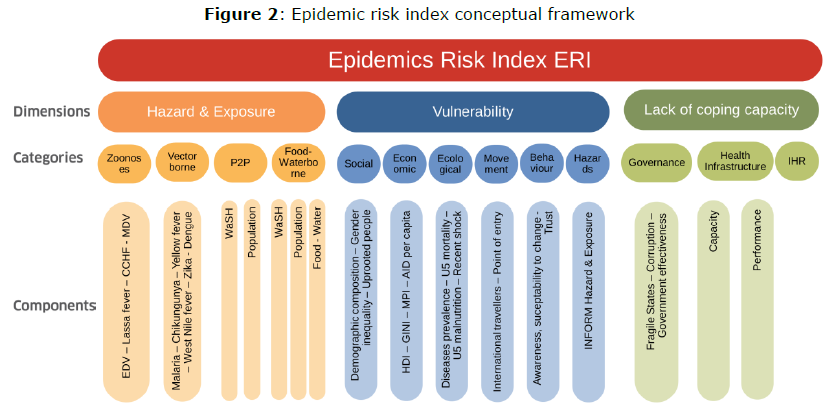
Data Collection

Data are collected through Python, except for health-related data (e.g. dengue):

|  |  |
| --- | --- |
| Data | Source |
| Time-series Temperature | https://www.accuweather.com/ |
| Dew Point | https://www.wunderground.com/ |
| Humidity | https://www.wunderground.com/ |
| Wind Speed | https://www.wunderground.com/ |
| Gust | https://www.wunderground.com/ |
| Pressure | https://www.wunderground.com/ |
| Precipitation Rate | https://www.wunderground.com/ |
| Accumulative Precipitation | https://www.wunderground.com/ |
| UV | https://www.wunderground.com/ |
| Solar Radiation | https://www.wunderground.com/ |
| time-series NDVI | https://modis.gsfc.nasa.gov/ |
| time-series EVI | https://modis.gsfc.nasa.gov/ |
| Tweet Count | twitter.com |
| Tweet Engagements (retweets, etc) | twitter.com |
| Tweet Sentiment (positive, negative, neutral) | twitter.com |
| Google trends | google.com |
| geospatial NDWI | https://landsat.gsfc.nasa.gov/ |
| geospatial NDBI | https://landsat.gsfc.nasa.gov/ |
| geospatial NDVI | https://landsat.gsfc.nasa.gov/ |
| geospatial Aerosol Index | https://landsat.gsfc.nasa.gov/ |
| Surface Temperature | https://landsat.gsfc.nasa.gov/ |
| AOI Polygon Geojson | https://boundingbox.klokantech.com/ |
| Hospitals and Clinics | https://www.openstreetmap.org |
| Dengue case count | https://doh.gov.ph/ |
| Dengue deaths | https://doh.gov.ph/ |

Risk-based Framework

To further strengthen the relevance of AEDES, one of the major enhancements is to integrate a risk assessment framework for dengue in regions and provinces of the Philippines using the INFORM Epidemic GRI Model. It has the potential to be adapted for evaluating dengue risk at both the national and local levels given that the model is applicable to several types of risks and epidemics.



Use-Cases

* Hazards – monitor progress of epidemic, generate alerts
* Vulnerabilities – prioritize areas with vulnerable groups, suggest demographic and geographic determinants of risk
* Coping Capacity – prioritize areas for emergency aid, recommend infrastructure investment

INFORM AEDES Data

The following datasets per use-case are aimed to be used in the risk-based assessment, subject to data availability and sustainability.

|  |  |  |
| --- | --- | --- |
| Hazards | Vulnerabilities | Coping Capacity |
| Dengue Case incidence  Flood Occurrence  Temperature  Precipitation  COVID-19 Incidence  Access to water  Access to sanitation | Population ages 0-20  Poverty Index  Population affected by natural disasters  Population previously infected by dengue  Mortality  Land-use types  Social listening  Primary and secondary schools  Philhealth coverage  Human mobility | Presence of health centers  Presence of hospitals  Number of health workers  Health expenditure  Vaccination coverage |

Project Plan

Expected Outcomes

|  |  |
| --- | --- |
| **OUTCOME** | **DESCRIPTION** |
| **Data Collection and Processing** | Database Management and Automated Data Ingest for Search Trends, Climate, Satellite, and Health Data which entails continuous research on alternate global open data sources. |
| **Data Analysis** | Incorporate Socio-Economic indicators using Dengue RISK INFORM in the predictive modeling and deploy to all regions. Enhancement of Dengue Case and Deaths Nowcasting. |
| **Product Development** | Redesign AEDES interface and functionalities which include information portals, publicly-accessible APIs, and near-real-time daily updates.  This will entail Dengue Trends Overview, Outbreak/Epidemic Monitoring (time-series projections, dengue hotspot map visualizations), At-Risk Community Assessment (risk ranking of regions and provinces, risk maps), and Actions and Recommendations. |

Risks

|  |  |  |  |
| --- | --- | --- | --- |
| **Risk**  **(description)** | **Level of risk**  **(high, low)** | **Probability**  **(high, low)** | **Action planned**  **to mitigate** |
| Data (Timeliness, Accuracy) | High | Low | Reliance on credible 3rd parties (NASA, Landsat, ESA) |
| LGU Buy-in | Medium | Medium | Lobbying local government partnerships through UN system (UNDP, WHO, UNICEF) |
| Dedicated staffing due to reliance on volunteers | Medium | Low | Dedicated staffing costs to hire full time dedicated staff as part of the project (50% of funding) |

Project Activities

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Month** | **Feature details** | **Status** | | **Result** |
| 1 month | API Development - Satellite (NDWI, NDVI, NDBI, Aerosol Index, Solar Irradiance/Surface Temperature) from Sentinel HUB, Landsat, Euro Data Cube | Refining | Automate data collection, repository, and feeder into application for remote sensing and dengue case count nowcasting/forecasting | |
| Socio-economic and weather data (hospitals, population, temperature, humidity, solar radiation, precipitation rate, cumulative precipitation) |
| 1 month | Risk INFORM Framework and Model Integration | Developing | Risk modeling incorporating hazard, vulnerability, and resilience indicators | |
| 0.5 months | Social Listening - Google Trends, Twitter count, engagements, sentiment | Refinement | Automated data gathering of social indicators | |
| 0.5 months | Dengue Trends Overview, Outbreak/Epidemic  Monitoring, At-Risk Community Assessment, Action and Recommendations | Refinement | Descriptive, Predictive, and Prescriptive  Modules | |
| 0.5 months | Dengue case and deaths nowcast enhancement | Refinement | Show results of prediction including more regions. | |
| 1.5 months | UI/UX Enhancement | Refinement | Mobile Responsive Web-App | |

Project Design

|  |  |
| --- | --- |
| **Milestone** | **Details** |
| Team Formation and Kickoff | Finalize team skill requirements  Role design  Promote job openings  Recruit team members  Onboarding and debriefing |
| Automate data gathering | Finalize automation requirements and python package design  Automation design  Remote Sensing  Google Trends  Weather  Data ingestion design - dengue cases and deaths  Implement automation  Remote Sensing  Google Trends  Weather  Implement data ingestion - dengue cases and deaths  Automation testing and evaluation, and finalizing data gathering python package |
| Enhance nowcasting models | Nowcasting design  Data gathering  dengue cases and deaths  google trends, twitter data  weather and satellite data  Nowcasting modeling  correlations and linear regression  time-series forecasting  AutoML  Nowcasting testing and evaluation  Nowcasting model deployment |
| Enhance mosquito hotspot detection model | Finalize hotspot detection requirements  Hotspot detection design  Data gathering  remote sensing  ground observations  Hotspot modeling through geospatial clustering with exogenous data  Hotspot model testing and evaluation  Hotspot model deployment |

|  |  |
| --- | --- |
| Web / mobile portal development and enhancement | Finalize web and mobile portal requirements  Web and mobile portal design and wireframes  Web and mobile portal development  Web and mobile portal testing and evaluation  Web and mobile portal go-live |
| Implement INFORM Risk Framework | Finalize risk requirements  Risk framework design  INFORM inputs data gathering  Implement INFORM Risk Framework  Risk Framework testing and evaluation  INFORM model deployment |
| Publication and dissemination | Finalize publication requirements  Publication plan and design  Manuscript drafting and development  Editing and feedback  Publication  Webinar and training |

Organizational Profile

The following resources are involved in this project.

|  |  |  |
| --- | --- | --- |
| **NAME** | **RESPONSIBILITIES** | **EMAIL** |
| **EMILY JO VIZMONTE,**  Project Coordinator and Research Lead | Manage and oversee the project implementation and provide oversight and direction to project activities. | emilyjo.vizmonte@cirrolytix.com |
| **XAVIER PUSPUS,**  Data Science and Machine Learning Lead | Prepare data for processing and analysis, leverage machine learning techniques for interpretation and insight generation | xavier.puspus@cirrolytix.com |
| **MARK NEIL PASCUAL,**  Platform Development and Engineering Lead | Develop the product design of the AEDES platform and manage the API infrastructure. | markneil.pascual@cirrolytix.com |

*\*\*\*End of this Document\*\*\**