**TEMPLATE 2: PRODUCT REQUIREMENTS DOCUMENT**

**Company Name: Cirrolytix Research Services - Project AEDES**

**PART 1:**

1. **Describe the primary use cases (real-world application) for your solution.** Include here both actual and potential use cases of your solution. Please provide data/results of any pilots. For each actual and potential use case, provide an explanation why you feel the technology is suitable for this use case.

|  | **Use case 1 (Real-world application)** | **Use case 2** | **Use case 3** |
| --- | --- | --- | --- |
| Brief Description  (200 characters limit) | Monitor for dengue panics online | Nowcast dengue cases and deaths | Identify mosquito hotpots on the ground |
| Actual or Potential | Actual | Actual | Actual |
| If available, summarise pilot results that confirm the use case (provide links to data and results):  (200 characters limit) | Google Trends identified 5 areas (as at August 2019) where dengue panics were >1000% of 2017 levels: Western Visayas, Zamboanga, Eastern Visayas, Bicol, and BARMM  https://dx.doi.org/10.2139/ssrn.3902598 | Model fitting showed nowcasting successfully predicted actual cases for PH National, Quezon City, Tacloban City, Iloilo City, and Cotabato City (R 0.75 - 0.9, R2 0.56 - 0.81)  https://dx.doi.org/10.2139/ssrn.3902598 | Mosquito hotspots were successfully visualized for Quezon City, Tacloban City, Iloilo City, and Cotabato City  https://dx.doi.org/10.2139/ssrn.3902598 |
| Why technology is suitable for this use case: (200 characters limit) | System data sources include aggregated Google Trends | Modeling uses google trends, precipitation, and temperature to successfully now cast dengue cases and deaths | Remote sensing indicators NDWI, FAPAR, and NDVI are able to identify likely mosquito habitats on the ground |

1. **What are the features within your solution that you will develop / refine / test over the duration of the investments (12 months)?**

| **Month** | **Feature details** | **Status (e.g. developing, refining)** | **Result (e.g. parts developed, feature tested and bugs fixed, level of accuracy etc)** |
| --- | --- | --- | --- |
| 3 months | API Development - Satellite (NDWI, NDVI, FAPAR) from Sentinel HUB, Landsat, Euro Data Cube  Socio-economic Data | Refining | Automated data collection, repository, and feeder into application for remote sensing |
| 2 months | Risk INFORM Framework and Model Integration | Developing | Risk modeling incorporating hazard, vulnerability, and resilience indicators |
| 2 months | Social Listening - Google Trends, Twitter Sentiments | Refinement | Automated data gathering of social indicators |
| 2 months | Dengue Trends Overview, Outbreak/Epidemic Monitoring, At-Risk Community Assessment, Action and Recommendations | Refinement | Descriptive, Predictive, and Prescriptive Modules |
| 2 months | Dengue case and deaths nowcast enhancement | Refinement | Show results of prediction including more regions. |
| 1 month | UI/UX Enhancement | Refinement | Mobile Responsive Web-App |

**PART 2:**

**Please choose the most relevant technology category (/ies) for your solution and complete questions on either**

**A) Software;**

**B) Hardware or**

**C) Content**

\*Make sure all relevant components are covered. You can provide answers in one, two or three categories.

\*Note that the questions for start-ups using data and / or blockchain technologies in section (A) Software and the questions for start-ups using drones technologies in section (B) Hardware; should only be answered if relevant to your solution.

**SECTION A: SOFTWARE**

1. **Prototyping: Describe the work you have done on proofs of concept, mockups, and user testing.** List the last three versions of your solution’s design. For each version, please provide 2 sentences that describe the reasons for the changes (including data) and outline the main changes made.

| **Version** | **Main characteristics and features** | **Main changes from previous version** | **Reasons for changes made**  (incl data collected/ feedback received/ refined understanding of user needs): |
| --- | --- | --- | --- |
| Original 1.0 | NASA Space Apps prototype | Covers: Quezon City, Iloilo City, Cotabato City, Tacloban City  Linear Regression Fitting  NDWI+NDVI, NDWI+FAPAR hotspot detection | Original design for hackathon |
| Version 1.1 | UX Enhancement - promotion model | Minor tweaks for UI development for user-friendly demos, added w3.css template and improved mapbox layers | Improvements to the hackathon prototype after the discussion with WHO Philippines and OpenStreetMaps |
| Version 1.2 | FTW Prototype | Added CALABARZON Cities (Tagaytay, Cavite, Batangas)  Added ARIMA and VAR models | For The Women Foundation Volunteers added new data to the repository |
| Version 2 | Publication Model | Added information page, enhanced mapbox layers | Improved basic functionalities as part of publishing the research paper on AEDES |
| Version 3 (in-progress) | Risk management model | Integration of INFORM epidemic risk framework and learnings from existing work on COVID-19 surveillance | Generation of risk maps and more intuitive risk recommendations |

1. **Before you developed your own, or selected an existing one, what other similar technologies did you consider using, improving or iterating to solve the problem you want to solve?** Outline what other technologies were considered to address the problem you identified and why were they rejected as unsuitable.

|  | **Technology description** (200 characters limit) | **Unsuitable because:** (500 characters limit) |
| --- | --- | --- |
| 1 | Google Maps | Cost tiering for G-Maps was prohibitive for prototyping testing |
| 2 | Bootstrap CSS | Overly bloated CSS framework for what was required. Simpler frameworks (e.g. W3) do the job without unnecessary code bloat |
| 3 | Django Framework (Python) | Python performance issues and unnecessary back-end for prototyping |
| 4 | Flask Framework (Python) | Python performance issues and unnecessary back-end for prototyping |
| 5 | Laravel Framework (PHP) | PHP only limited to serving back-end data, Laravel adds code bloat without adding needed functionality |
| 6 | JQuery Framework (Javascript) | Outdated conventions already superceded by ES5/6 on vanilla JS |
| 7 | Angular Framework (Javascript) | Unnecessary code bloat |
| 8 | React Framework (Javascript) | Unnecessary code bloat |
| 9 | Plotly Charting Library (Javascript) | Visualizations are too heavy |
| 10 | Android Studio / Java Development | Native mobile app creates unnecessary redundancy in code |
| 11 | IoS/Kotlin/Objective-C Development | Native mobile app creates unnecessary redundancy in code |

1. Please select the license type that your solutions are currently registered under.

☐ AGPL

☐ LPGL

☐ BSD

☑ **Other equivalent (please specify): MIT License**

☐ Not currently on an open source license

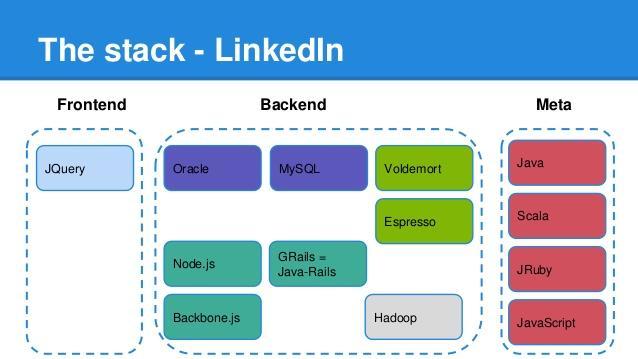
1. **If currently not on an open source license, are you willing to release your tech solution under an AGPL, LGPL, BSD or equivalent open-source license** (please refer to the Funding Agreement to consider the obligations for open source licensing and investment by the Innovation Fund):

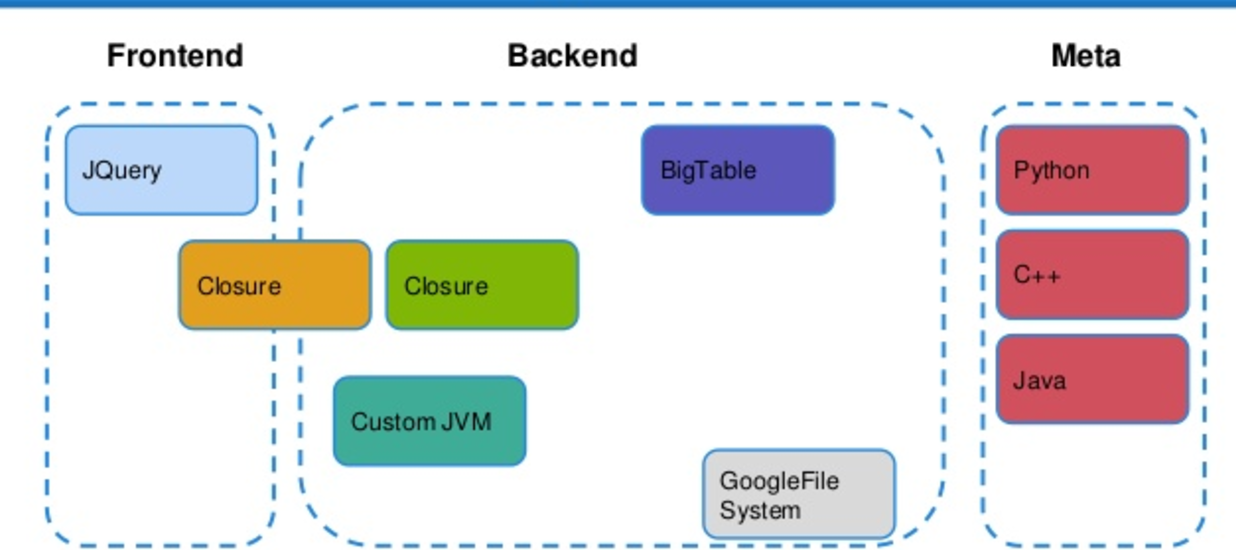
☑ Yes

☐ No

1. **Please insert at the end of this template (Annex 1) a diagram of your current tech stack including all components and subsystems, showing their linkages and interactions. A tech stack is a combination of software products and programming languages used. For this diagram, include: both client side and server side documentation if relevant, tools used for front-end and back-end systems (including databases), and any meta-layers**. **If you are using machine learning, please describe your pipeline.**

**Please find here two examples of tech stacks that you can use as templates:**





* **In addition, please provide written documentation explaining the diagram provided, outlining each element included in the diagram, its role and its status of development (complete, under development, expected date of completion etc). Please clarify if these components are currently open source or proprietary, and in the case of the latter if you intend to put them on an open source license. Include links or documents outlining process documentation, system documentation, user documentation, and any instructions on API usages, if relevant.** (5,000 characters limit)Technology Stack*Front-end Stack*
* W3-CSS - lightweight CSS framework for web-based and mobile-responsive front-end
* AJAX - Javascript asynchronous calls to back-end for data updates
* ChartJS - chart data visualization library
* MapboxJS - free tier of popular mapping library

*Back-end Stack*

* QGIS - open-source satellite API for remote sensing data extraction and analysis
* PHP - page routing and serving data access requests
* Python - perform machine learning (Scikit-Learn, STATS Models) and data analysis (Pandas)
* PostgreSQL - store and aggregate data
* Jupyter - data analysis and machine learning environment for Python

*Data Sources*

* NASA GIBS WMS - dynamically expose and visualize various satellite layers and indicators
* Sentinel HUB - access to and visualization of Sentinel satellite data
* Google Trends - access and extraction of public search interest
* DOH Epi Bureau - provision data on dengue cases and deaths
* DOST-PAGASA - provision data on average daily temperature and precipitation

High-level data architecture*Sentinel 2 and Landsat 8 (NASA GIBS WMS/Sentinel HUB)*

* Normalized Difference Water Index (NDWI), Fraction of Absorbed Photosynthetic Active Radiation (FAPAR), and Normalized Difference Vegetation Index (NDVI) extracted using QGIS
* Overlap of indicators provides data for mosquito hotspots which are ingested into Mapbox for visualization

*DOST-PAGASA*

* Provision CSV data on average daily temperature and precipitation

*Google Trends*

* Search interest on ‘dengue’, ‘dengue symptoms’, ‘dengue medicine’, and ‘dengue cure’
* Top search trends are visualized using ChartJS

*DOH Epibureau*

* Provision data on dengue case counts and deaths by region

*Nowcasting*

* Precipitation, temperature, and dengue search data were combined into a linear model to generate the nowcasts
* Nowcast data is visualized via Chart JS

1. **Where do you host the repositories of your solution:**
   1. ☑ **GitHub**
   2. **Bitbucket**
   3. **Other** (please specify)**:**
2. **Share here all GitHub / Bitbucket (or other) repositories for all components of your solution:** (1,000 characters limit)

<https://github.com/Cirrolytix/aedes_dpg>

<https://github.com/docligot/aedesproject>

<https://aedesproject.org/privacy>

If your repositories are private and in Github or Bitbucket **please make sure to give access to our user before you send this form back to us: @unicefinnovation or** [**venturefund@unicef.org**](mailto:venturefund@unicef.org) (we assure you that the information in your repositories will only be used to evaluate funding your solution and will not be shared).

1. **Please share the github / bitbucket (or other) handles of the code developers within your team:** (1,000 characters limit)

<https://github.com/docligot>  
<https://github.com/rachemelendres>

<https://github.com/metromark>

<https://github.com/fstayco>

<https://github.com/cricketsoong>

1. **How many external contributors have there been to your code**? (1,000 characters limit)

Apart from the original AEDES team, enhancements were contributed by members of the FTW Foundation team.

1. **Please share the stack overflow profile (or equivalent) of the code developers within your team (if available)?** (1,000 characters limit)

<https://stackoverflow.com/users/6609543/mark-toledo>

1. **What processes and tools do you leverage for quality assurance and testing?** (1,000 characters limit)
   1. Lighthouse Tool using Chrome DevTools for web accessibility [performance](https://github.com/Cirrolytix/aedes_dpg/blob/main/Lighthouse%20Chrome/AEDES%20Web%20Accessibility.pdf)
2. **What version control system do you use?**

Subversion:

☑ Git: <https://github.com/Cirrolytix/aedes_dpg>

Other:

1. **Sustainability: What is the ongoing plan for maintenance, updates, and support of your software?** Describe what the needs will be for the next two years and how these will be met.(1,000 characters limit)  
     
   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)
* 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)

1. **If your start-up is using data as a central part of the solution:**
   1. **Describe your data sets (i.e. size, features, method of collection, labels) If applies, please also include the description of the test and training data sets.** (1,000 characters limit)

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

* + 1. **Google Trends Data**

The raw values, ranging between 0-100, represent search interest as 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 the regional level using *dengue*, *dengue fever, dengue symptoms, dengue medicine, and dengue cure*

* + 1. **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.

* + 1. **Weather Data**

Average rainfall and temperature were sourced out from the available weather stations of the DOST-PAG-ASA in the targeted areas. Data is available via an official request addressed to the Climatology and Agrometeorology Division (CAD) of DOST-PAGASA with a corresponding paid service fee.

* + 1. **Satellite Data**

Process Sentinel2 band satellite data using QGIS 3.4 to determine stagnant water locations in the specified city. The generated output is a CSV file which contains the coordinates of these potential Dengue hotspots computed from the FAPAR and NDWI of the satellite image.

* 1. **Describe how you already collect or plan to collect data and with which technologies.** (1,000 characters limit)

We have developed data extraction manuals for maintaining and updating the datasets through these link: <https://github.com/Cirrolytix/aedes_dpg/wiki/D.-Explore-Updating-Datasets>

* Satellite Data: Satellite imaging data from [Sentinel Online Copernicus](https://sentinel.esa.int/web/sentinel/sentinel-data-access) using QGIS
* Local Weather Data:
  + - * Climate data from [DOST-PAGASA](http://bagong.pagasa.dost.gov.ph/climate/climatological-normals)
      * Alternate weather indicators: Land Surface Temperature and Precipitation both of which are available from Terra/Aqua and Global Precipitation Measurement (GPM)

<https://gpm.nasa.gov/data>

<https://lpdaacsvc.cr.usgs.gov/appeears/products>

<https://lpdaac.usgs.gov/tools/appeears/>

* Google Data: [Search trends for 'dengue' and related terms](https://trends.google.com/trends/explore?date=today%205-y&geo=PH&q=dengue) - automate extraction
* Disease Surveillance Data: Regional cases and deaths data from [Department of Health](https://doh.gov.ph/statistics)

* 1. **Describe who is the owner of the data and what licenses is it under.** (1,000 characters limit)
     1. **Satellite Data -** [**Landsat 8 (NASA)**](https://www.usgs.gov/land-resources/nli/landsat/landsat-8) **and** [**Sentinel-2 Copernicus (ESA)**](https://sentinel.esa.int/web/sentinel/sentinel-data-access) **- Open Data**
     2. **Local Weather Data -** [**DOST-PAGASA**](http://bagong.pagasa.dost.gov.ph/climate/climatological-normals) **- Paid service**
     3. **Google Search Trends - Google** [**https://policies.google.com/terms**](https://policies.google.com/terms) **- Open Data**
     4. **Disease Surveillance Data - Department of Health - Open Government Data via Freedom of Information Bill**

* 1. **If you use personally identifiable indicators, please describe how you secure the privacy of individuals.** (1,000 characters limit)

We use aggregated data only from legitimate sources and third-party paid service providers.

1. **If your start-up is using machine learning or A.I:**
   1. **Describe your models and why you settled on using them. (E.g. we use Logistic Regression because it is easy to interpret and we don't have enough data to use more sophisticated methods).** (1,000 characters limit)For dengue nowcast model fitting, we performed a multiple linear regression using temperature, precipitation, and each of the dengue search and monthly coding as fixed effects.

Training and Test Set

Using the lagged factors and dummy variables as predictors for dengue cases, we split the data into training and test sets. The training set covers all months between 2015 to 2017 while the test set for all monthly values of 2018.

Feature Selection

Select only those that have a significant relationship with the target variable. Only significant p-values are selected as features for statistical modelling.

Backward Stepwise Regression

With backward stepwise regression, the model starts will all the candidate features and gradually eliminates the least significant predictor at a time until all the remaining predictors have p-values smaller than the set threshold. The process generated a list of models with all possible combinations of predictors.

* 1. **How do you evaluate your models / methods? (i.e. accuracy, F1 score, Regression coefficient)** (1,000 characters limit)

Model Performance Evaluation

Three sets of models which passed the diagnostic tests and have the highest adjusted R-square values were selected and visualized in the information portal of the dashboard.

* 1. **Describe your roadmap to further improving your model and how you plan to maintain it and keep it up-to-date** (1,000 characters limit)Whilst linear model performed satisfactorily during prototyping exercise, we have plans to test other models for nowcasting. Version 1.2 by FTW featured time-series analyses such as ARIMA and VAR which we hope to continue testing. Additionally other models such as Support Vector Regressions will be tested for applicability.   
       
     Since new data is added from sources every day, model decay will be monitored on a monthly basis with a new review and recalibration triggered whenever model fit goes below R 0.71 / R2 0.5 or 12 months, whichever comes first
  2. **Describe the limitations of your methodology: Where does your approach fail? (e.g. your predictions are biased because you do not have enough data for a certain class)** (1,000 characters limit)Model fails to account for location based factors and population - both of which are expected to be addressed by the INFORM framework once incorporated.

1. **If your start-up is using blockchain:**
   1. **What blockchain platform are you using? (ie. Bitcoin, Ethereum, Hyperledger, etc). Please explain your choice and clarify whether it is a private or public chain.** (1,000 characters limit)

* 1. **Is your product deployed? If so, where? (ie. private node, mainnet, test net)** (1,000 characters limit)

* 1. **What consensus algorithm does the blockchain you’re using use? (ie. Proof of Work, Proof of Stake, Delegated Proof of Stake)** (1,000 characters limit)

* 1. **What language are you using to create your smart contracts? (ie. Solidity, Lvy, Viper).**  (1,000 characters limit)

* 1. **How do you go about testing your smart contracts?** (1,000 characters limit)

* 1. **Is personal/biometric data being stored in the blockchain? If yes, how are you ensuring the privacy of that data?** (1,000 characters limit)

**SECTION B: HARDWARE (leave blank, if no hardware-related components to your project)**

1. Please select the license type that your solution is registered under.

☐ CERN

☐ Other equivalent:

☐ Not currently on an open source license

1. **If currently not on an open source license, are you willing to release your tech solution under a CERN or equivalent open-source license** (please refer to the Funding Agreement to consider the obligations for open source licensing and investment by the Innovation Fund):

☐ Yes

☐ No

1. **In the space provided at the end of this template (Annex 2), provide a detailed diagram of the hardware you are developing.** Include block diagrams of overall system architecture, hardware schematics, images of your system, and detailed drawings and descriptions (this can include CAD or 3D diagrams, electronics layouts, rough BOM, and links to relevant documentation).
2. **In addition, please provide written documentation explaining the diagram provided, outlining each element included in the diagram, its role and its status of development (complete, under development, expected date of completion etc). Please clarify if these components are currently open source or proprietary, and in the case of the latter if you intend to put them on an open source license. Include links or documents outlining process documentation, system documentation, user documentation.** (5,000 characters limit)

1. **Please provide a component list with supplier/ SKU numbers and origin (where do you source each component from).** Be sure to list every component displayed in the exploded view in Annex 2.

1. How many units of your product are currently available?

1. **a) How much does** **each unit cost to produce?**

**b) How does the cost change as the number of units increase?**

1. **Where do you manufacture your product?**

1. **If you are using drones in your solution please specify:**
2. **What hardware components are you buying off-the-shelf and which are you producing yourself? Please specify for each of them if they are available in your country or if you import them.** (1,000 characters limit)

1. **What is the overall cost of your hardware device? (Including off-the shelf and self-produced elements)** (1,000 characters limit)

1. **How many flights did you make to date (approx.)?**
2. **What is the average distance you are covering with your drone? With what payload?**
3. **What was the longest distance you have been able to cover so far? What was your payload with that distance?**
4. **What is the maximum payload your drone can carry? What is the distance that the drone can fly with that payload?**
5. **Did your drone crash; if so, why? What did you change as a lesson learned?** (1,000 characters limit)

1. **Can you install a tracker (e.g. ADS-B) to allow the surveillance of your flight?** (1,000 characters limit)

1. **What fail-safes do you use?** (1,000 characters limit)

1. **Do you have permission to fly your drone in your country? Did you get permission for the tests you are planning on doing? If not, what is the timeline and process in getting that?** (1,000 characters limit)

**SECTION C: CONTENT**

1. **Please provide an overview of the new content you have created or are planning to create.** Provide a description that includes but is not limited to: subject, audience, language, format, medium. (1,000 characters limit)

1. **Who else is developing similar or identical content? How is your new content different from the existing one? What is the gap you are attempting to address?** (1,000 characters limit)

1. **List the learning objectives you are planning to meet through the content**. (1,000 characters limit)

1. **Please insert on the last page of this template (Annex 3), a content/curriculum map diagram (max 1 page), including the general organization of the content, rough relational hierarchy.** **Please clarify if any these components are currently open source or proprietary, and in the case of the latter if you intend to put them on an open source license.** Be sure to give details at the lowest level of granularity of your content.
2. **Please provide a sample of the content you have created (at the end of this template Annex 4) and/or plan to create (copy/paste at the end of this form in the appropriate place).** The sample should cover one piece, ie one video, one pdf, one set of exercises etc.
3. **Provide a short explanation that shows how this sample relates to the overall content.** Describe the context of the example content within your larger set of content. (1,000 characters limit)

1. **Please describe how the content will be delivered to the user (face-to-face facilitation, self-paced, etc).** (1,000 characters limit)

1. **Please explain where the content will be hosted and how the user will access the content** (LMS, CMS, etc.). (1,000 characters limit)

1. **Provide the number of content objects (including the breakdown of media types)…**

a) ...currently offer:

-audio (e.g. mp3)

-video (e.g. mp4)

-rich text (pdf, html, other markup language)

-pure HTML5 frontend apps (embeddable in iframe)

b) propose to create:

-audio (e.g. mp3)

-video (e.g. mp4)

-rich text (pdf, html, other markup language)

-pure HTML5 frontend apps (embeddable in iframe)

1. **Please select the license type that your solution is registered under.**

☐ CC-BY

☐ CC-BY-SA

☐ CC-BY-NC

☐ CC-BY-ND

☐ CC-BY-NC-SA

☐ CC-BY-NC-ND

☐ Other equivalent (please specify):

☐ Not currently on an open source license

1. **If currently not on an open source license, are you willing to release your tech solution under a CC-BY or CC-BY-SA or equivalent open-source license** (please refer to the Funding Agreement to consider the obligations for open sourcelicensing and investment by the Innovation Fund):

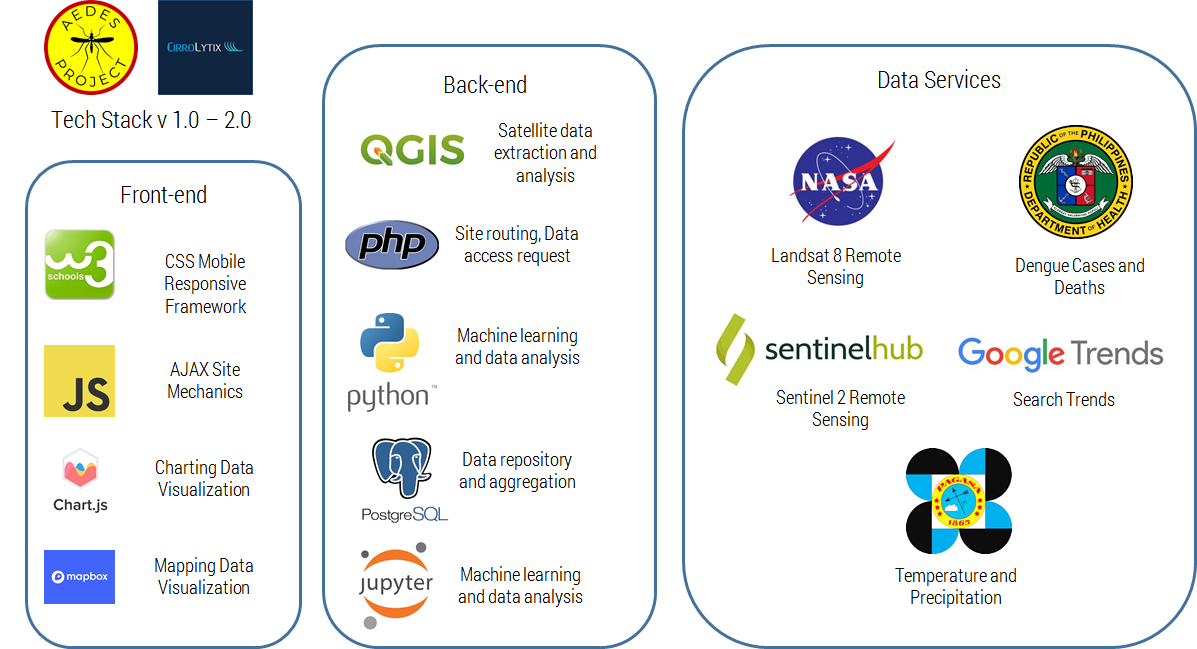
☐ Yes

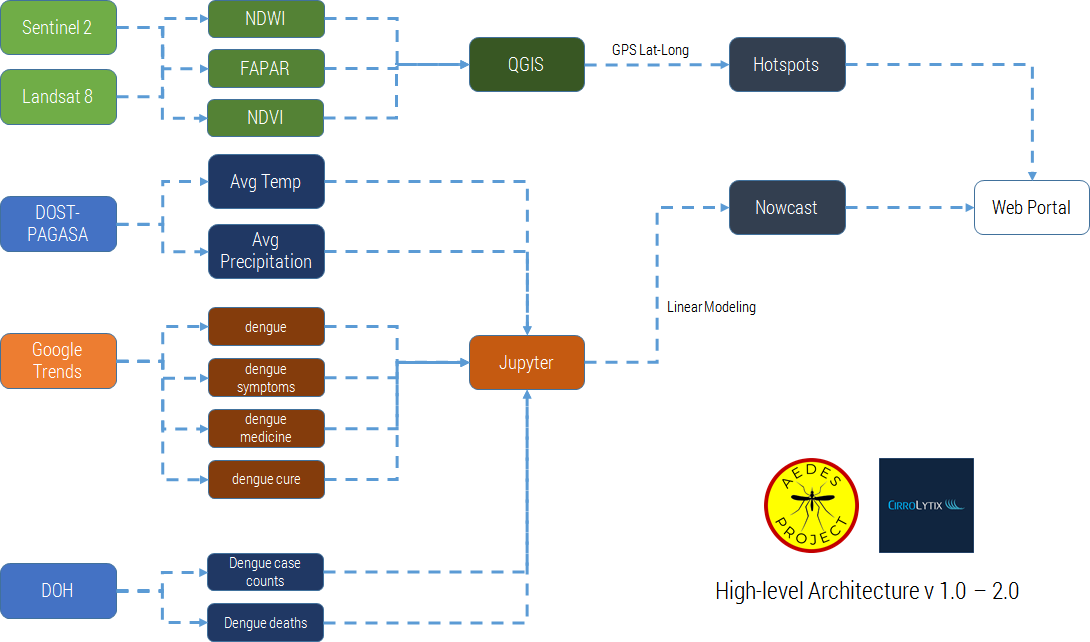
☐ No

1. **Sustainability:** How will content be updated based on user needs, feedback and data collected? (1,000 characters limit)

**ANNEX 1: SECTION A. QUESTION 5**

***PLEASE INSERT HERE: Annex 1:* a diagram of your current tech stack including all components and subsystems, showing their linkages and interactions. A tech stack is a combination of software products and programming languages used. For this diagram, include: both client side and server side documentation if relevant, tools used for front-end and back-end systems (including databases), and any meta-layers**.

****

****

**ANNEX 2: Section B, Question 3**

**Insert here: detailed diagram of the hardware you are developing.** Include block diagrams of overall system architecture, hardware schematics, images of your system, and detailed drawings and descriptions (this can include CAD or 3D diagrams, electronics layouts, rough BOM, and links to relevant documentation).

**ANNEX 3: SECTION C. QUESTION 4**

***PLEASE INSERT HERE: Annex 3:* a content/curriculum map diagram (max 1 page), including the general organization of the content, rough relational hierarchy.** Be sure to give details at the lowest level of granularity of your content.

**ANNEX 4: SECTION C. QUESTION 5**

***PLEASE INSERT HERE: Annex 4:* a sample of the content you have created (at the end of this template Annex 4) and/or plan to create (copy/paste at the end of this form in the appropriate place).** The sample should cover one piece, ie one video, one pdf, one set of exercises etc.