

Request for Proposals - VIVID for Malaysia
Questions and Answers (10.22.2025)

Topic	Sub-topic	Questions	Answer
Automated monitoring and detection of advertising or commercial content	Data type	Which types of data does the software need to process (text/ image/ video) ?	The data types to be processed mainly include text and images, with a small portion of short videos from Facebook.
	Data source	Which sources does the software need to process (Social media/ other websites)?	Data sources include Facebook, Instagram, brand websites, and e-commerce platforms.
	Scalability	How many data sources will be crawled?	The sources are not limited, they will depend on the technical capability of the crawler.
		What is the frequency of the crawling module: Daily/ Weekly?	It is expected that data collection will occur on a daily basis
Automated pre-qualification of the legal or transgressive nature of the content identified based on the detection criteria	Data sample	Can you define the output labels (violation categories, criteria)?	The main classification criterion will be whether the content constitutes a violation or non-violation. Subcategories will specify the type of violation, based on the International Code of Marketing Breast-milk Substitutes and the national code of Malaysia.
		Can you provide a sample of the data (including input data and output results)?	Example (Vietnamese post): Input data: https://www.facebook.com/share/p/1B7YHRf1Zx/ The system will crawl and store the link and its content in the database. Output data: Violation status: Violated Type of violation (based on keyword or image): Promote milk formula for children under 24 months / Gifts, coupons, vouchers A similar tool for Vietnam can be found on: https://code.corporateaccountabilitytool.org/vietnam
	Evaluation method	How do we define the evaluation method?	The system's performance will be evaluated through expert testing after each model development stage, accompanied by a result report.
	Accuracy Rate	What is the expected Accuracy Rate?	The expected accuracy rate is 70–80%.
	Training/ hosting infrastructure	What is the expected hosting infrastructure for AI models (Self-hosting/ Cloud API)?	Both are acceptable; however, it should be declared in the proposal.
Localization (consisting of translating and adapting the content to Malaysian cultural specificities)		Please clarify the action of adapting the content to Malaysian cultural specificities.	"Adapting" means aligning with the national framework in Bahasa Malaysia, with updates to the homepage (introduction text, visuals, and layout) accordingly.
Continuous Learning Phase		Please clarify the term "Continuous Learning" in the Specific Activities table.	The model is expected to learn and improve based on the data provided, with the capacity to add more data sources over time.
		What is the duration of the Continuous Learning phase?	This phase will continue until the end of the contract period and potentially beyond, focusing on testing and improving accuracy.
Maintenance Phase		Please define the scope of work of the Maintenance phase. Does this phase introduce new test cases for model evaluation?	The maintenance phase aims to ensure the smooth operation of both the crawler and the model.