This project is funded by the European Union

13 June 2023

Energy Efficiency Compliant Product 2018 – EEPLIANT3 [GA 832558]

Call for Tender

Tender Specifications

in connection with the procurement of services for developing a web crawler prototype or a minimum viable product (MVP) under the EU-funded EEPLIANT3 Concerted Action.

1. CONTEXT AND SCOPE OF THE PROCUREMENT

The EEPLIANT3 project is an EU-funded concerted action (https://eepliant.eu/index.php/new-about-eepliant/about-eepliant3) that aims to support the implementation of EU Energy labelling and Ecodesign regulations and improve the energy efficiency of products placed on the Single Market. Besides executing targeted market controls, the 29 participating Market Surveillance Authorities (MSAs) also seek to co-create methodologies and systems to harmonise their work and achieve efficiencies.

One of the objectives of the EEPLIANT3 project is to develop digital tools to support MSAs in enforcing EU product energy efficiency legislation. The procurement is **to design and build a working beta prototype or, at best, an MVP of a pro-active e-Commerce web crawler** by the end of November 2023.

1.1 Why we want to develop a web crawler prototype to support online market controls relating to consumer products subject to Energy labelling and Ecodesign EU legislations.

National market surveillance authorities (MSAs) are responsible for verifying whether products sold in the EU, online and through physical shops, satisfy the minimum requirements set out in Ecodesign and Energy labelling regulations.

Part of these requirements concerns the obligation of internet retailers/web shops or other online service providers to display the electronic energy label and product information sheet in accordance with the provisions of EU regulation. The Ecodesign directives set minimum performance standards for a range of products.

When performing market surveillance checks online, MSA inspectors will first select a number of retailer websites in their national market depending on their own available resources, the market structure and based on previously established priority criteria, and then, depending on the energy-related product that is targeted, e.g. televisions, will visit each selected web shop to verify whether the product pages for televisions comply with the EU requirements, i.e., that all the mandatory/essential product-related information and energy label are present, accurate and displayed correctly. This operation also involves the manual extraction/recording of all relevant information.

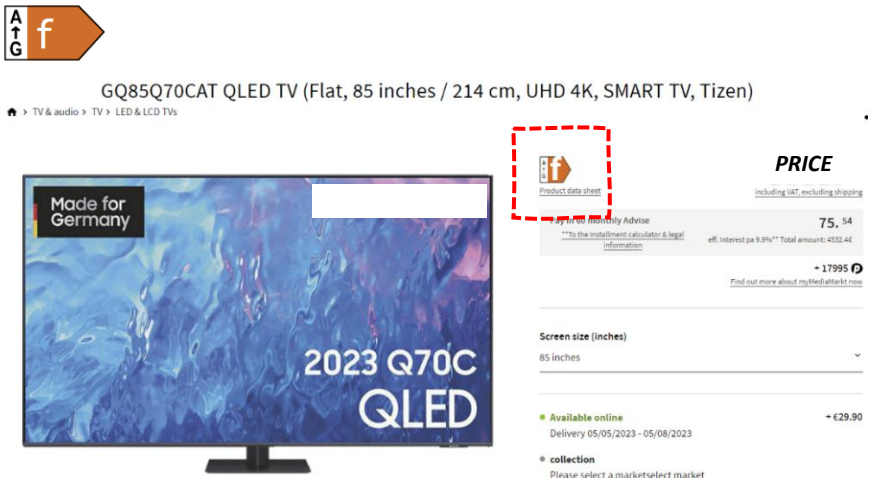
This manual verification and data collection requires a great deal of time and effort for one to check a significant number of product pages and record the information. With this prototype, we aim to start building a sustainable solution to automate the data extraction (first, from a small number of pre-selected websites) in order to create efficiencies for the inspectors.

2. DEFINITIONS

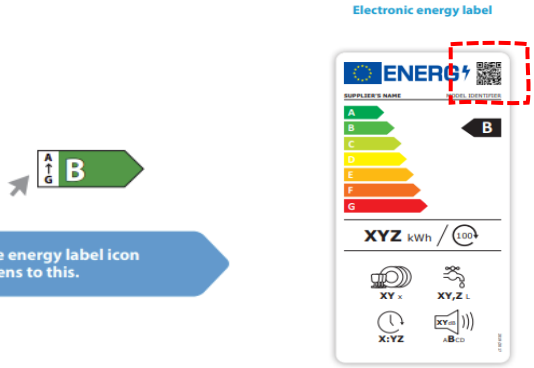
General note:

The energy label and product information sheet must be displayed on retailer webpages where the price is displayed, and the relevant products can be ordered. In general, there are two main display options for the energy label and product information sheet on a webpage which will be discussed below.

A-G energy label: An energy label image as a nested display present on a webpage that sells any of the 15 product types which require an energy label according to the EU regulation:



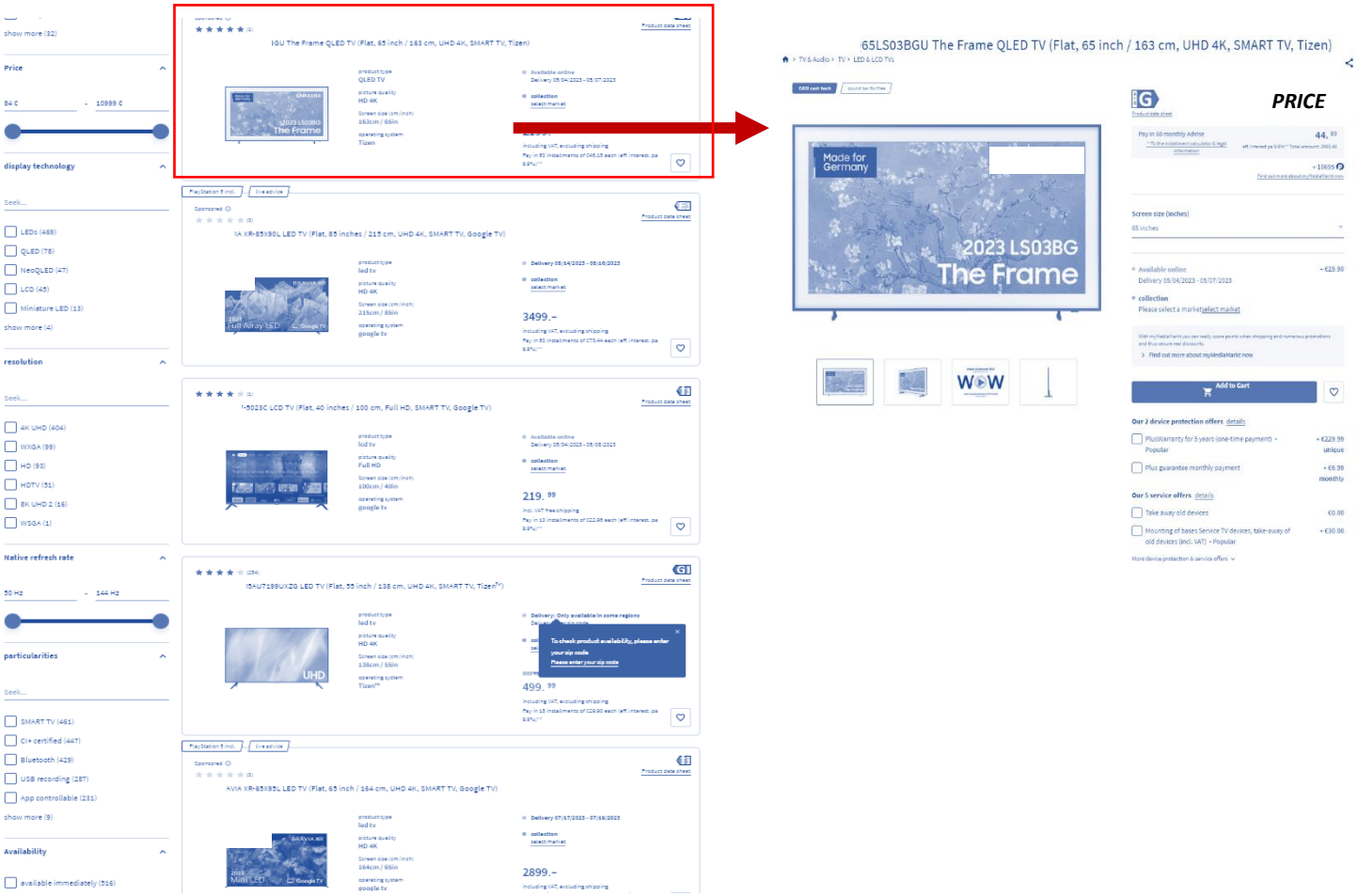
EU Energy Label: The EU Energy Label displayed (by pop up, new tab or new page) after clicking on the A-G energy label image or displayed directly as a full-size picture/image on a product page. Besides all other essential information, the energy label also includes a QR code in the upper right corner (see image below). This QR code provides a direct link to official (non-commercial) information for a model, as they have been inserted by manufacturer into the European Product Registry for Energy Labelling (EPREL) - for EPREL see below.



The QR code

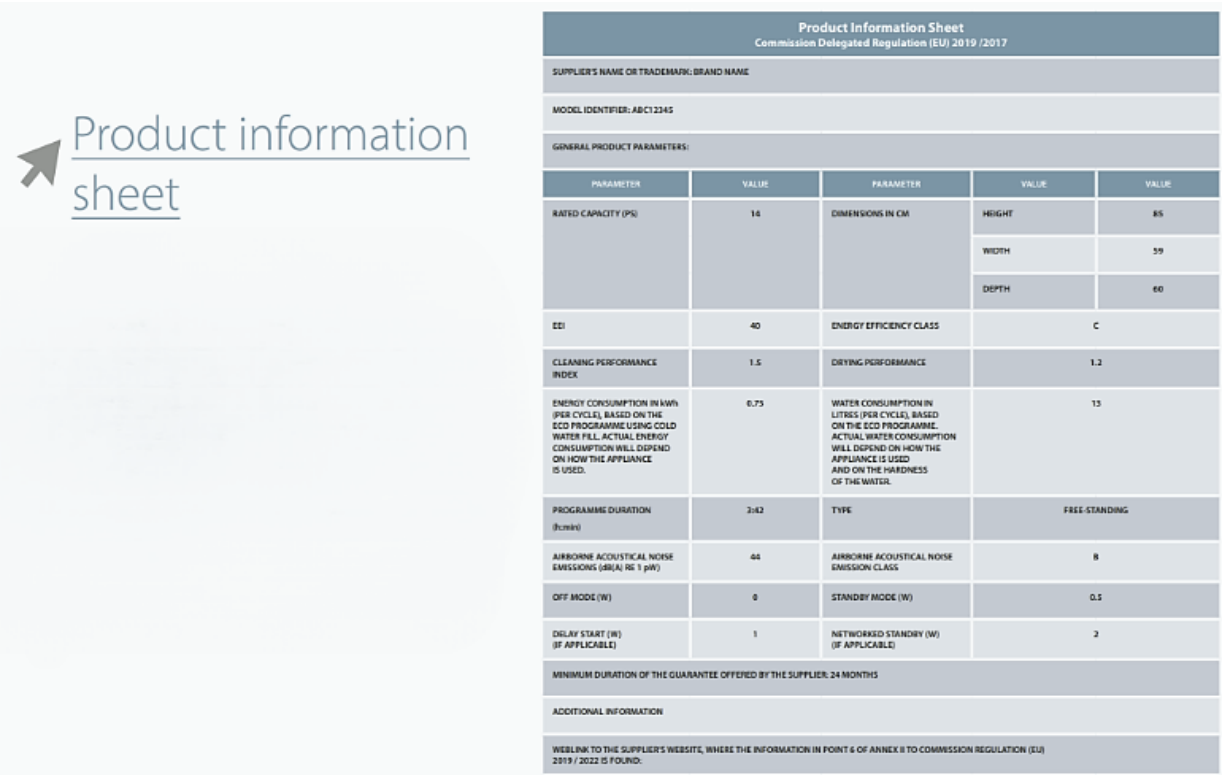
Product page: A web page as part of the structure/web taxonomy of a retailer website that markets and sells a model of a specific product category, presenting information about the product like price, size consumption etc. (typically, 1 product model = 1 product page):

Product page



Product information

sheet: Products for sale online must be accompanied also by a product information sheet. A product information sheet is an electronic document or page that provides customers with product information, including specific energy performance and environmental related information about a product. A product page should either display directly the full product information sheet or provide a link to an electronic version (as provided by the manufacturer or the importer).



Target application/ software: The custom software of a web crawler prototype or MVP which shall be the primary product of the upcoming service contract with the selected software developer under the EEPLIANT3 concerted action.

European Product Registry for Energy Labelling (EPREL): EPREL is an online database created by the European Commission to provide consumers and dealers with additional information about specific energy-related products and to ensure a more effective compliance control by market surveillance authorities. As of 1 January 2019, suppliers must register their products in EPREL. In May 2022, the database was launched for public access and consultation. The public part of EPREL provides the label data and the product information sheet for each registered product which can be viewed and downloaded through a QR code shown on the energy label.

(https://commission.europa.eu/energy-climate-change-environment/standards-tools-and-labels/products-labelling-rules-and-requirements/energy-label-and-ecodesign/product-database_en)

3. DESCRIPTION OF THE SERVICE

The service concerns the design and development of a working prototype or MVP of a pro-active e-Commerce web crawler. The key functional and technical requirements have been specified also through feedback collected during two online info sessions on 25 and 27 April 2023 as part of an open market consultation and following the publication of a prior information notice.

3.1 Overview of key functional and technical specifications and system requirements

Description of the target software/application:	A working prototype or MVP of a web crawler/web scraper that will be able to scrape product information/data from at least three retailer websites in one EU Member State market.
Product type in scope	Televisions
Product sub-types	LCD, LED, OLED, QLED, Plasma
Min. # of websites ("root URLs") to be crawled	5
Minimum functional requirements	1. GUI-based operation ready to scale up across variable cross-platform frameworks (desktop, mobile) 2. Multi-user operation through individual password-protected user accounts; Implementation of two roles: administrator and user 3. Ability to identify all the relevant product pages to be crawled in the web taxonomy of each selected website 4. Scrape/Extract data relevant to the selected product (televisions) from the product pages of the pre-selected web shops 5. Ability to launch a search on demand and also automatically/on schedule 6. Store and index all the collected data in a database 7. Timestamp data collection 8. Normalise/Clean the extracted data 9. Ability to export structured data to Excel and csv 10. Ability to import data to Microsoft Power BI 11. Identify if the A-G energy label or the full-sized energy label is present (on a True/False condition) and, in case of a nested display, identify whether clicking on the A-G label launches the full EU Energy Label

	12. Read the QR code displayed on the EU energy label and then download the label information and all the product data from the public part of European Product Registry for Energy Labelling (EPREL) 13. Take a snapshot of the energy label for export 14. Recognise/Identify if a URL link to the product information sheet exists or not on every product webpage, and if so, copy the URL and download that sheet as an image 15. Anti-bot bypass technology [IF NEEDED for this iteration, depending on the retailer websites that will be selected, see also below Section 9: Financial Offer]
Target information for extraction [to be refined together with the preferred bidder]	– brand – model identifier – product-type/sub product type – product identifier – panel technology (e.g., LED) – price – model year – energy efficiency class at Standard Dynamic Range (SDR) – energy efficiency class with high dynamic range (HDR) – scale of energy efficiency classes from A to G – On-Mode Power Consumption at Standard Dynamic Range (SDR) – On-mode power consumption with high dynamic range (HDR) – On-mode energy consumption at standard dynamic range (SDR) – Power consumption in standby mode – Power consumption in networked standby mode – screen size in cm – screen size in inches – screen resolution – energy label (nested and/or full size) – data on the product information sheet URLs of all product pages crawled

Anticipated system components	1. Web crawling engine 2. Database 3. Web client (if needed) 4. Data cleansing module 5. Data calculation module (second priority)
Mandatory technical requirements relating to the proposed technology/ programming language	There are no prior restrictions with regards to the technology and programming language that can be used for developing the prototype. However, the proposed technology must ensure **scalability and adaptability, i.e., the prototype must be built in such a way that is scalable and adaptable, including the possibility of covering in the future more EU markets, more retailer websites and product types, and adding more functionalities, deeper automation and machine learning.**
Mandatory technical requirements relating to the cloud hosting service or local/on-site solutions.	We recognize that both cloud-based and on-site (localised) solutions have their own strengths and weaknesses, and we are seeking a comprehensive analysis that will help us to make an informed decision about which option is best suited to our needs. We request the service supplier to offer both technologies and include an analysis with a detailed comparison of the features and benefits of each option, as well as any potential drawbacks or risks associated with each approach. The analysis shall include a comparison of costs for both technologies and advise accordingly. Requested minimum requirements: - Compliance with GDPR (General Data Protection Regulation) - Servers used are hosted in the EU - ETC
Mandatory requirements regarding ownership rights, copyrights, and IPR in relation to the new custom software/target application	PROSAFE and the EEPLIANT3 beneficiaries shall retain **full and exclusive** joint rights, title or interest in and to the target application including but not limited to codes, patents, inventions, or other intellectual property developed or created in the course of this project, together with the right to market, sell, and distribute the target application.
Minimum system requirements	– Multi-user operation through individual password-protected user accounts; Implementation of two roles: administrator and user – Simultaneous access to and real-time use of the system by multiple users – Ability to back up files on demand

Language version of the operating system	English
Maintenance requirements after the handover of the prototype to PROSAFE	None. The software code should be stored in GitHub, ready to be restored for continuing development in the future.
Deadline for the delivery of the working prototype/MVP to PROSAFE	**24 November 2023**
KPIs / Key user acceptance criteria, after UAT	**Target rate** [indicative, TBD together with the selected contractor]
Target crawling speed and frequency	TBD = non-essential requirement for this first iteration, but **the system must be as optimised as possible** taking account of the number of pages to be crawled, the volume of the data to be scraped, and the benchmarks of other similar solutions
Classification accuracy	0.60
Precision: i) Accurate data points extracted divided by the total number of data points extracted. ii) Ratio of product pages crawled to actual existing/relevant product pages per website	0.60 0.75
True/False precision for energy label	0.70
NB: These KPIs are rough indications and additional KPIs will be defined by the team of subject matter experts together with the selected developer	

3.2 Description of the minimum desired features as user stories

In this table we capture the desired software features in the form of user stories. It should be read together with Section 3.1. Each user story has a priority class. This has been specified through a prior MoSCoW analysis, where:

M = Must have: Non-negotiable product need that are mandatory and must be delivered

S = Should have: Important feature that creates significant value, but which can be excluded from this iteration, if it is difficult to deliver within time and cost constraints

C = Could have: Nice to have feature but with limited impact/added value

W = Won't have: Feature that is not a priority for this iteration or cannot be delivered within the given time and cost constraints

DESCRIPTION OF USER STORIES		PRIORITY
User Story #1	As a market surveillance inspector, I want to be able to automatically extract product-related information from the targeted retailer websites, so that I can check compliance with ED/EL regulations faster and more efficiently.	Must have
User Story #2	As an inspector, I want the web crawler to flag/indicate somehow the product pages which potentially do not display an energy label, so that I can have a first signal of potential non-compliance with (EU) 2019/2013.	Must have
User Story #3	As an inspector, I want the web crawler to be able to identify when essential product information is missing on a product page, so that I can quickly have a first indication of potential non-compliance with ED/EL obligations.	Must have
User Story #4	As an inspector, I want the web crawler prototype to store, index and clean the extracted information on a secure database, so that I can access and work on it in real-time and simultaneously with my peers.	Must have
User Story #5	As an inspector, I want to the web crawler to be able to read the QR code on the energy label on each product page, so that it can then download the label and all the available product data from the EPREL database.	Must have
User Story #6	As an inspector, I want to be able to use the collected data, so that I can analyse product availability in the market, do some sort of statistical analysis, and inform my organisation's future market surveillance plans and priority-setting.	Must have
User Story #7	As an inspector, I want to be able to access the URL of the crawled product pages via a user interface, so that I can quickly compare and verify the information extracted with the actual information shown on the product page of the retailer website.	Must have
User Story #8	As an inspector, I want the web crawler to have a desktop user interface that I can navigate easily, so that I can launch a new search and access all the information that were extracted and stored in the database through a personal password-protected user account.	Must have
User Story #9	As an inspector, I want to be able to program the web crawler, so that it can automatically crawl websites periodically.	Could have

User Story #10	As a programmer/software developer, I want to be able to easily read and modify the code of the web crawler prototype, so that I can scale up and develop more modules and functionalities, including artificial intelligence (AI), and adapt to future needs. See section 3.4 below.	Must have
User Story #11	As a programmer/software developer, I want the web crawler prototype/MVP to be built in such a way, so that it will be ready to scale up in the future and integrate further AI features and data analytics modules.	Must have
User Story #12	As an inspector, I want to be able to import the extracted data to MS Power BI, so that I can create statistics, a dashboard, and manipulate them.	Must have
User Story #13	As an inspector, I want to be able to export the data stored in the database to Excel, so that I further process them.	Must have
User Story #14	As an inspector, I want the web crawler to be able to avoid bot detection, so that it can scrape a website with a bot detection system in place that can block web scraping.	Could have
User Story #15	As an inspector, I want the web crawler to be able to automatically calculate the Energy Efficiency index and check if the energy label/class is the correct one, in order to save time.	Must have
User Story #16	As a system administrator, I want to be able to easily create new user accounts or delete existing ones, so that I can give or suspend access to the system according to needs.	Must have
User Story #17	As an inspector, I want the system to be able to back up the extracted data that are saved in the database, so that I can restore them and avoid data loss.	Must have
User Story #18	As a software developer and if the database will be stored on the cloud, I want to be able to export and save it locally, so that I can access the data at any time offline/outside the cloud.	Should have
User Story #19	As a user/inspector, I want to be able to customize the search parameters (filters, specifying keywords, or selecting data sources)	Must have
User Story #20	As a user/inspector, I want the web crawler to be able to collect the data of the product information sheet to calculate benchmarks and process consistency checks	Must have

3.3 Non-functional requirements

The MVP shall be developed with the following non-functional requirements in mind:

Scalability: The WebCrawler prototype/MVP will be able to handle an increasing volume of requests as the user and website base grows. KPIs could include the ability to add or remove resources quickly, the ability to handle spikes in traffic, and the ability to scale horizontally or vertically.

Performance: The WebCrawler prototype/MVP will be able to process a large number of requests within a reasonable amount of time. The performance shall be measured E2E which includes the API and its databases. KPIs could include response time, throughput, and capacity utilization.

Reliability: The WebCrawler prototype/MVP will be able to operate consistently and reliably over time, with minimal downtime or failures. KPIs could include availability, uptime, and mean time between failures (MTBF).

Security: The WebCrawler prototype/MVP will be designed and implemented with appropriate security measures to protect against unauthorized access, data breaches, and other security risks. KPIs could include the number of security incidents, the frequency of security audits, and the level of compliance with industry standards and regulations.

Maintainability: The WebCrawler prototype/MVP will be designed and implemented with ease of maintenance and updates in mind. KPIs could include the time required to perform updates or maintenance tasks, the number of bugs or defects identified and resolved, and the level of automation in the testing and deployment processes.

3.4. Code Documentation Requirements

The WebCrawler's code shall be thoroughly documented in English language. The documentation shall include clear and concise explanations of the code's purpose, design decisions, data structures, and algorithms used. It shall also include instructions on how to install, configure, and deploy the WebCrawler on different environments. The documentation shall be kept up-to-date and always made available to PROSAFE.

4. PROJECT TEAM and WAY OF WORK

The selected service supplier will work closely with an external IT consultant in the Role of a Product Owner - contracted by EEPLIANT3, the management team of the project led by PROSAFE, and a dedicated group of subject-matter experts from participating market surveillance authorities, forming a cross-functional team.

We would like to request that the supplier adopts an agile way of working and ask the supplier to work closely with a team of subject matter experts to deliver product increments over the lifetime of the project. We are looking for a supplier who shares our commitment to collaboration, communication, and responsiveness and delivers product increments at regular intervals throughout the project lifecycle, with each increment providing incremental value to the end user.

The approach should be based on regular feedback loops and collaboration with the team, enabling flexibility, adaptability, and responsiveness to change. We expect the supplier to provide regular updates on progress, respond promptly to feedback, and adapt plans and timelines as needed to ensure that the project stays on track. We believe that an agile approach will help us to achieve our goals more efficiently and effectively, while at the same time minimizing product risks.

5. PLACE OF PERFORMANCE OF THE CONTRACT

Main site or place of performance: Contractor's premises. No travelling is foreseen for this service.

6. TERM OF SERVICE

The service duration is expected to be from mid-July to end of November 2023.

7. EXCLUSION CRITERIA

Bidders found to be in any of the following exclusion conditions will be rejected and excluded from the evaluation procedure:

a) they are bankrupt or being wound up, are having their affairs administered by the courts, have entered into an arrangement with creditors, have suspended business activities, are the subject of proceedings concerning those matters, or are in any analogous situation arising from a similar procedure provided for in national legislation or regulations;

b) they have been convicted of an offence concerning their professional conduct by a judgment which has the force of res judicata;

c) they have been guilty of grave professional misconduct;

d) they have not fulfilled obligations relating to the payment of social security contributions or the payment of taxes in accordance with the legal provisions of the country in which they are established or with those of the country of the contracting authority or those of the country where the contract is to be performed;

e) they have been the subject of a judgment which has the force of res judicata for fraud, corruption, involvement in a criminal organisation or any other illegal activity detrimental to the European union's financial interests;

f) following another procurement procedure or grant award procedure financed by the Community budget, they have been declared to be in serious breach of contract for failure to comply with their contractual obligations.

Tenderers are asked to provide a Declaration on Honour (see Appendix 1) stating that they are not in one of the situations giving rise to exclusion from the procedure, as listed above.

Acceptance of the standard terms and conditions for PROSAFE contracts

The submission of an offer implies that bidders:

- Accept PROSAFE's "General Conditions for Tenders" (Appendix 2).
- Accept that CINEA (EU Agency managing the Horizon 2020 programme), the European Commission, the European Court of Auditors, and OLAF (European Anti-Fraud Office) have the right to carry out checks, reviews and audits on contractors and subcontractors during or after the implementation of the EEPLIANT3 Concerted Action.

8. EVALUATION CRITERIA

Bids will be assessed according to the following evaluation criteria:

CRITERION 1:	**Experience in/Capacity for developing web crawlers/web scrapers (10 points)**
Evidence to submit:	Description of the web crawler or web crawlers developed by the bidder or by software developers/programmers currently employed by the bidder who will be involved in the development of the web crawler prototype for EEPLIANT3. Description should include scope, technology(-ies)/ programming language used per system, time required until deployment.
CRITERION 2:	**Description of the proposed solution (20 points)**
Evidence to submit:	Description of the proposed custom software solution, including its **technical specifications**, the **proposed solution for the IT Infrastructure and database** to be used **(cloud hosting service, local infrastructure, or combination?)**, the **project roadmap with a breakdown of milestones and tasks (from July to Nov 2023)**, and **how the proposed solution matches/satisfies overall the functional, technical and other requirements** as described in Section 3 above.
CRITERION 3:	**Identification of risks and mitigations (10 points)**
Evidence to submit:	A list of risks and proposed mitigation measures in relation to the development and deployment of the custom software, from inception to delivery.

Each criterion will be awarded a score from 0 (zero) up to a maximum number of points as indicated above. **The total maximum number of points is 40.**

9. FINANCIAL OFFER

Bidders are requested to submit a financial offer for the delivery the requested service as described in Section 3 above. For the purposes of this procurement, bidders should submit an offer for each of the following two hypothetical configurations:

SCENARIO A (Baseline scenario):

- Number of retailer websites ("root URLs") to be crawled: 5

- Type of retailer website to be crawled: Multinational chains of stores selling consumer electronics and/or large (= with significant market share) national/market-specific online retailers
- Number of markets/countries: 1 = Ireland
- **Provide two prices:** One with anti-bot bypass technology and a data calculation module (see User Story #15), and another without these two.

SCENARIO B:

- Number of retailer websites ("root URLs") to be crawled: 20
- Type of retailer website to be crawled: Multinational chains of stores selling consumer electronics and/or large (= with significant market share) national/market-specific online retailers
- Number of markets/countries: 2 = Ireland, Germany (10 websites per country)
- **Provide two prices:** One with anti-bot bypass technology and a calculation module (see User Story #15), and another without these two.

The financial offer should indicate:

a) The estimated total price **for each Scenario**, including all charges/expenses for the development of the target application as described in Section 3 above, **excluding potential costs relating to cloud hosting** (if any, see also below). **If the solution includes a cloud hosting service, the service to be used will be selected together with the preferred bidder. Costs relating to cloud hosting will be covered by PROSAFE.**

b) The hourly rate to be charged.

All quoted prices should be expressed in EUROS, **with zero-rated VAT**.

PROSAFE is VAT registered as taxable person established in Belgium with VAT number BE 0809.226.854 (Belgian Chamber of Commerce 809 226 854). All invoices shall mention the BE VAT number and be issued with zero VAT, making reference to the reverse charge mechanism according to Articles 44 and 196 of the VAT Directive 112/2006.

Terms of offer must be valid for acceptance (or negotiation) for at least 3 months from submission. PROSAFE reserves the right to negotiate with one or more shortlisted bidders before taking a decision on the placing of a contract. The final specifications of the service will be determined in the contracting phase together with the selected supplier.

Invoicing will be discussed and agreed before the placing of the contract.

10. TENDER DOCUMENTATION

To be considered complete and go through the evaluation process, tenders should include:

a) A signed Declaration on Honour (see Appendix 1) sent in digital scanned copy via email, as specified in Section 7 above (the selected bidder will be asked to provide an original copy via post during contracting)

b) Evidence addressing all the Evaluation Criteria as described in Section 8 above.

c) A financial offer as per Section 9.

11. ASSESSMENT OF TENDERS

Bids complying with the exclusion criteria as defined in Section 6 above will be evaluated based on the following weighting:

- **Evaluation criteria: 70%**
- **Financial offer: 30%**

The selection procedure will be as follows:

1. Screening of tenders for compliance with the exclusion criteria. Non-compliant tenders are rejected.
2. Assessment of the qualifying bids based on the weighted **evaluation criteria.**
3. Possible direct negotiations with up to three shortlisted bidders.
4. Bidders are informed on the results of the tender evaluation.
5. Award of a fixed-term service contract to the bidder who demonstrates best value for money.

Bidders may be invited to provide additional information to clarify the already presented services or where a clerical error occurred provided that the principles of transparency and equal and fair competition are respected.

12. SUBMISSION

The deadline for submission of offers is **03 July 2023, 17:00 CEST**.

Bids must be sent to PROSAFE in hardcopy (Avenue des Arts/Kunstlaan 41, 2nd floor, 1040 Brussels, Belgium) **AND** via email to ioana@prosafe.org **and** kyriakos@prosafe.org with the subject header 'EEPLIANT3 WebCrawler Tender'. Hardcopies must be received at latest by 07 July 2023 – stamp date being the proof that they were sent no later than 03 July 2023.

Tenders received after this deadline will be rejected.

13. QUESTIONS

The deadline for submission of questions is **28 June 2023**.

Questions may be sent to Kyriakos Papazoglou, Programme Manager, at Kyriakos@prosafe.org. The responses that PROSAFE will provide to all questions received will be published here:

https://eepliant.eu/index.php/calls-for-tenders?id=181-procurement-of-service-development-of-a-webcrawler-prototype and https://prosafe.org/index.php/en/news-events/calls-for-tender?id=853

In respect of equal treatment and confidentiality, confidential or commercially sensitive information will not be published in the Q&A. The Q&A will be published in an anonymised way without specific details about suppliers or their solutions.

14. DATA PROTECTION

Bidders must ensure compliance with the applicable data protection rules at national and EU level including Regulation (EU) 2016/679 (General Data Protection Regulation) on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)[1].

Additional background information

Additional information about the EEPLIANT3 Concerted Action can be found here: https://eepliant.eu/index.php/new-about-eepliant/about-eepliant3

Additional information about Work Package (WP) 2 of EEPLIANT3 on the development of Digital Tools for market surveillance on Energy labelling and Ecodesign is available here: https://eepliant.eu/index.php/horizontal-themes/it-tools

With best regards,

Ioana Sandu

Executive Director

Appendices

Number	Title
1	Declaration of Honour
2	PROSAFE's General Conditions for Tenders

[1] https://eur-lex.europa.eu/eli/reg/2016/679/oj

Commission or any other body of the European Union, who are not responsible for any use that may be made of the information it contains.