

# Q&A 1

## Public European tender

### Prototype Automatic Arabic Dialect Recognition (A-ADR)

TN 562514

#### Questions

| No. | Document  | Chapter  | Paragraph or clause   | Page no. | Question   | Answer IND  |
|-----|---|--|---|----------|--|---|
| 1   | Tendering Guidelines + Annex 2 (Price) + Annex 5                | Tendering Guidelines – Definitions; 2.2.2 Objective of the Assignment; Annex 2 – Sub-award criterion Price | Tendering Guidelines – 'Further agreement' (Definitions table)<br><br>Tendering Guidelines – Objective 1: Determination of parameters (months 1–4)<br><br>Annex 2 – 'Sub-award criterion Price' + 'Please note: minimum/maximum price...' | 6; 10; 5 | The tender documents establish that the determination of permissible training datasets is a central deliverable of Objective 1 (months 1–4), and that the scope of dialect coverage for model development directly depends on the outcome of that determination. As a consequence, the total training workload for Objectives 2 and 3 cannot be reliably estimated prior to contract signature. Given this dependency chain, we would like to ask whether the Contracting Authority considers the following pricing approach valid and compliant with the tender requirements:<br><b>Objective 1</b> priced as a fixed-scope, fixed-price phase covering the determination of parameters as defined in section 2.2.2.<br><b>Objectives 2–5</b> priced as a variable component, with the final price determined after completion of Objective 1, capped within the stated bid validity range of €600,000–€1,200,000 excl. VAT, and formalized through a Further Agreement as defined in the tender's definitions. | We are looking for a fixed price offer for the entire project (all objectives). The functional requirements are given in Annex 1 2.1. We estimate that the selection of datasets will mainly affect the granularity and performance that can be achieved, not the workload.   |
| 2   | Annex Requirements A-ADR IND                                    | 3 Security requirements (Security-by-design) / 4 Project Governance  | 3.1 Information security – Security-by-design (development/deployment environment constraints)<br><br>4.1 Meetings and reports (collaboration with IND)   | 4; 5     | Where will the prototype be deployed during development — IND's infrastructure, contractor's infrastructure, or a neutral environment? What are the network/connectivity constraints?  | The prototype should be developed on the Contractor's infrastructure. The IND will not require access during development.   |
| 3   | Annex Requirements A-ADR IND                                    | 1 2 Functional requirements; 3 Security requirements   | 2.1 Prototype (functional context)<br><br>3.1 Information security – Security-by-design (controls impacting on-prem/cloud choices and constraints)  | 3; 4     | What are the compute resource constraints (GPU availability, on-prem vs. cloud)? Is there a preferred cloud provider or is on-prem mandatory?  | For the prototype we have no such specific constraints in mind. This is up to Contractor.   |
| 4   | Annex Requirements A-ADR IND; Annex 5 Concept overeenkomst      | 1 Annex 1 – 2 Functional requirements; Annex 5 – Article 2 (Object of the Agreement)                       | Annex 1 – 2.1 Prototype (processing/usage context)<br><br>Annex 5 – 2.1 (completion criteria tied to deliverables/validation context)   | 3; 3     | What is the expected processing throughput — batch processing only, or must the system support near-real-time analysis?  | We do not know the answer to this for the Final Solution. For the Prototype, batch processing is sufficient.  |
| 5   | Tendering guidelines_A-ADR tool; Annex 1 Requirements A-ADR IND | Tendering Guidelines – 2.2.2 Objective of the Assignment; Annex 1 – 3.3 Privacy and data protection        | Tendering Guidelines – Objective 1: 'determining which speech datasets are allowed...' (months 1–4)<br><br>Annex 1 – 3.3 Privacy and data protection (dataset/data-handling constraints)  | 8; 5     | Is there any specific datasets IND has already pre-approved or is actively considering? Are any existing licensed corpora (e.g., MGB, MADAR, ADI) already in scope?  | We have various options in scope, but we are awaiting legal advice on some details. Compliance with GDPR, copyright and license agreements are of course imperative. The NFI report (see Tendering guidelines, 2.1.3) includes an appendix and a discussion of Arabic datasets that have often been used in ADR research. See also answer to Q16. |

|    |   |  |  |             |  |  |
|----|---|--|--|-------------|--|--|
| 6  | Tendering guidelines_A-ADR tool; Annex 1 Requirements A-ADR IND                               | Tendering Guidelines – Definitions; Annex 1 – 3.3 Privacy and data protection                            | Tendering Guidelines – 'Processor agreement' (Definitions table)<br><br>Annex 1 – 3.3 Privacy and data protection (access/logging/retention/organisational measures context)   | 6; 5        | Who owns the data pipeline infrastructure? Will the contractor process data on IND systems, or will IND transfer data to contractor environments under the processor agreement?                                  | During the assignment all data, code etc. should be processed and stored by the Contractor. Nothing will be run on IND systems. The IND dataset of 293 anonymized use case recordings will be made available under an agreement.   |
| 7  | Annex 1 Requirements A-ADR IND; Annex 5 Concept overeenkomst; Tendering guidelines_A-ADR tool | Annex 1 – 1 General requirements; Annex 5 – Article 2; Tendering Guidelines – 2.2.8 Open source          | Annex 1 – requirement on open-source tool / license determined by Contracting Party<br><br>Annex 5 – 2.1 (open-source license; dataset/licensing constraints)<br><br>Tendering Guidelines – 2.2.8 Open source (license to be determined by Contracting Authority)                | 3; 3; 12    | What open-source license type is expected for the prototype — MIT, Apache 2.0, GPL, or other? Any forbidden third-party components (e.g., models with non-permissive licenses)?                                  | The type of license will be determined as part of Objective 1 (months 1-4). Any third-party components to be used are to be discussed in that period as well, including a check of their license conditions and intellectual property. Any third-party components should not block the transferability of the prototype to another party, should we decide to have a Final Solution developed. |
| 8  | Annex 1 Requirements A-ADR IND  | 2 Functional requirements  | 2.1 Prototype – broad level dialect classification (broad classes / 'best matching language model')  | 3           | What are the exact target dialect classes at broad level? The document lists Maghrebi/Levantine/Mesopotamian/Gulf/Egyptian — is this the confirmed taxonomy, or subject to change during Objective 1?            | This taxonomy makes sense, but could be subject to change during Objective 1.  |
| 9  | Annex 1 Requirements A-ADR IND  | 2 Functional requirements  | 2.1 Prototype – fine-grained dialect classification (sub-varieties)  | 3           | For the fine-grained level, is Syria-specific dialect distinction (Northern vs. Southern Levantine) a hard requirement or dependent on data availability only?   | A more fine-grained distinction can be EITHER towards country level OR it can focus on language varieties within the Syrian geographical area (e.g. Northern versus Southern Levantine), depending on availability of suitable data.   |
| 10 | Tendering guidelines_A-ADR tool; Annex 1 Requirements A-ADR IND                               | Tendering Guidelines – 2.2.1 Scope of the European tender; Annex 1 – 2 Functional requirements           | Tendering Guidelines – scope (prototype development/testing; mock-up context)<br><br>Annex 1 – functional requirements (no explicit integration requirement → clarification requested)   | 10; 3       | Must the prototype integrate with any existing IND systems (e.g., case management, identity systems) even in prototype stage?  | No, there will be no such integration at this stage. The IND is currently not envisaged as provider of the Final Solution.   |
| 11 | Annex 5 Concept overeenkomst; Tendering guidelines_A-ADR tool                                 | Annex 5 – Article 2 (Object of the Agreement); Tendering Guidelines – 2.2.1 Scope of the European tender | Annex 5 – 2.1 (validation reports; reliability; factors affecting reliability; presentation for lay users)<br><br>Tendering Guidelines – scope: 'comprehensive reports on performance' + mock-up/uncertainty communication context   | 3; 10       | What is the required output format for the dialect classification result — probability distribution, likelihood ratio, log-likelihood ratio, or verbal conclusion scale (e.g., as used in forensic linguistics)? | This is not fixed yet. As part of SAC 3 tenderers are invited to submit (1) a proposal for performance measures that are helpful to provide sufficient clarity on the performance of the Prototype, and (2) your ideas on how to present the output to a potential user of a Final Solution (e.g. an asylum officer).  |
| 12 | Annex 5 Concept overeenkomst; Annex 2 Sub award criteria; Annex 1 Requirements A-ADR IND      | Annex 5 – Article 2; Annex 2 – Sub-award criteria; Annex 1 – Prototype/validation requirements           | Annex 5 – 2.1 (completion tied to deliverables; no explicit accuracy threshold stated)<br><br>Annex 2 – quality criteria descriptors (how quality is assessed)<br><br>Annex 1 – requirements for prototype/validation (no explicit 'minimum accuracy' → clarification requested) | 3; 2-4; 3-4 | Is there a defined minimum accuracy threshold or performance benchmark the prototype must meet to be considered successful?  | No. Whether the results are deemed useful enough also depends on its sensitivity to various factors. A-ADR will probably work relatively well in certain conditions but less well in others. A clear picture of the conditions in which it works well - and how well - is needed before deciding on the next step (implementation or not).   |

|    |   |  |  |  |  |   |
|----|---|--|--|--|--|---|
| 13 | Annex 1 Requirements A-ADR IND; Tendering guidelines_A-ADR tool                 | Annex 1 – 3.3 Privacy and data protection; Tendering Guidelines – Definitions                  | Annex 1 – GDPR requirements (personal data processing, controls)<br><br>Tendering Guidelines – 'Processor agreement' (Definitions table)   | 5; 6                                     | Must all data processing occur within the Netherlands/EU, or are other jurisdictions acceptable for development/training compute?  | All data processing must occur within the EER or one of the countries outside the EU that have an adequate level of data protection, as determined by the EU on the basis of article 45 of Regulation (EU) 2016/679<br><br>Please refer to:<br><a href="https://commission.europa.eu/law/law-topic/data-protection/international-dimension-data-protection/adequacy-decisions_en">https://commission.europa.eu/law/law-topic/data-protection/international-dimension-data-protection/adequacy-decisions_en</a>  |
| 14 | Tendering guidelines_A-ADR tool; Form A Uniform European Tender documentEA_ESPD | Tendering Guidelines – Supporting documents / verification stage; ESPD/UEA – Exclusion grounds | Tendering Guidelines – 'Certificate of good conduct in tendering' (Justis for NL; eCertis for foreign equivalents; ≤2 years freshness)<br><br>Form A (ESPD/UEA) – exclusion grounds declarations (context: what is declared vs. what may be evidenced later) | 14; (Form A – exclusion grounds section) | Certificate of good conduct in tendering: please confirm<br>(a) at which stage it will be requested (submitted with the bid vs. requested only from the intended awardee upon verification),<br>(b) whether it must cover the legal entity only or also persons with powers of representation, decision or control,<br>(c) whether a foreign equivalent identified via eCertis is acceptable and, specifically, whether a 'Criminal records certificate' would be accepted for foreign companies, and if so whether it should be issued for the legal entity and/or for the company's director(s)/persons with representation/decision/control powers,<br>(d) the required form (original paper vs electronically signed) and any language/translation requirements, and<br>(e) the required 'freshness' interpretation (not older than 2 years as of the submission deadline).  | a) Only from the intended awardee upon verification<br>b) The economic operator itself or any person who is a member of its administrative, management or supervisory body or has powers of representation decision or control therein (we refer to the document ESPD on Tendered)<br>c) We refer to what is stated in eCertis<br>d) Both are accepted. In Dutch or English<br>e) Not older than 2 years as of the submission of your tender  |
| 15 |   |  |  |  | <b>Open-source licensing and intellectual property rights</b><br>Regarding Requirement 2 (Annex 1), which states that "The AADR prototype must be developed as an entirely open-sourced tool under an appropriate license":<br>Can the Contractor retain ownership of background intellectual property (pre-existing technologies, algorithms, and methodologies) developed prior to this project? Specifically, we seek clarification on whether:<br>•The Contractor can continue to use and commercialize the underlying technology and core algorithms for other clients<br>•Only the specific prototype implementation and project-specific adaptations need to be open-sourced<br>•The Contractor retains rights to reuse training methodologies and technical approaches in future commercial projects<br>Without clarity on background IP protection, it is difficult to assess project feasibility and commercial viability. | We aim to have a prototype developed as an open source tool for a number of reasons, including transparency, transferability to another party (in case we wish to have a Final Solution developed), and to avoid vendor-lock-in at this point. Any third-party elements, choice of data or background IP should not hinder the transferability. Proprietary non-transferable elements (data and code) that are critical to the system's working should therefore be avoided. Possible non-critical third-party or background IP elements (e.g. for score calibration) will be discussed and identified as part of Objective 1. The IND will not claim IP. The Contractor would be free to reuse and commercialize technology, algorithms and methodologies for other clients. |

|    |                        |          |   |    |  |   |
|----|------------------------|----------|---|----|--|---|
| 16 |                        |          |   |    | <p><b>Dataset approval process and timeline</b></p> <p>Requirements 12 and 13 (Annex 1) state that only IND-approved datasets may be used, with datasets to be agreed upon during the first four months after contract signature.</p> <p>Can the Contractor receive preliminary approval or guidance on acceptable datasets and their legal status BEFORE contract signature? We need to understand:</p> <ul style="list-style-type: none"> <li>•What categories of publicly available datasets are pre-approved for use?</li> <li>•IND's criteria for dataset approval (privacy, copyright, licensing requirements)</li> <li>•Whether the Contractor can propose a preliminary dataset list for IND's legal review during the tender phase</li> </ul> <p>Without advance certainty about usable training data, we cannot reliably commit to delivering the required performance metrics and timeline.</p> | <p>See also answer to Q5. We understand that not knowing what data may be used forms a challenge when preparing bids. In addition to Annex 1 point 12, we can offer the following principles:</p> <ol style="list-style-type: none"> <li>1. For licensed data, e.g. those offered by LDC, the license terms should allow for future commercial use (i.e. use is not restricted to research purposes only).</li> <li>2. Datascraping from YouTube (necessary to access ADI17 dataset) by the Contractor will likely not be allowed.</li> <li>3. Use of proprietary data for training will not be allowed since this will result in vendor lock-in. Proprietary data could perhaps be used for testing only.</li> <li>4. The available metadata should allow for a preliminary evaluation of the usefulness (especially geographical spread of origins).</li> <li>5. Whether datasets are representative/good enough for the intended purpose (cf. AI Act Article 10) is to be established empirically through testing the models trained with these data. Contractor can propose datasets for review, but final decisions can only be taken during Objective 1.</li> </ol> |
| 17 |                        |          |   |    | <p><b>Risk allocation for dataset restrictions</b></p> <p>Given Requirements 12-13 (Annex 1), what provisions exist if IND's legal approval process significantly restricts available training datasets after contract signature?</p> <p>Specifically:</p> <ul style="list-style-type: none"> <li>•If approved datasets prove insufficient to meet performance requirements (e.g., dialect coverage, accuracy thresholds), can project scope or success criteria be adjusted accordingly?</li> <li>•Will the Contractor be held liable for non-performance if dataset restrictions make technical requirements unachievable?</li> <li>•Can the contract include provisions for scope adjustment based on final approved dataset capabilities?</li> </ul> <p>Clear risk allocation is essential for responsible bidding on this fixed-scope contract.</p>   | <p>The assignment does not include performance requirements in the sense of accuracy thresholds. The scope is primarily to provide sufficient clarity on the performance of a Prototype (reliability of the output and factors affecting this reliability) to allow further decision making on possible operational deployment, while compliance to all EU regulations are key. Contractor will not be held liable for any lack of performance. Scope adjustment is possible during the assignment.</p>   |
| 18 | Tendering guidelines_4 | 4,3      | Supporting documents to be submitted                | 14 | <p>Can eCertis be used by a UK-based company? An initial search for equivalent documents for the certificate of good conduct and the tax authorities certificate did not yield any results.</p>  | <p>Following documents apply:</p> <p>Certified Copy of Register Information (Certificate of Good Standing): To be requested from Companies House. This document confirms that the company is officially registered and not in liquidation.</p> <p>Basic DBS Check (Disclosure and Barring Service): evidence of the absence of relevant convictions.</p> <p>Letter from HMRC: To demonstrate compliance with tax and social security obligations (as a substitute for the tax declaration often requested alongside the Certificate of Good Standing).</p>  |
| 19 | Annex 1                | 3.1 / 33 | Information security - Use of Accredited QA systems | 4  | <p>Is our understanding that specific QA/QMS systems must be used in order to be compliant with the EU AI act? If so are there any recommendations, so that we can factor this into the costings?</p>  | <p>A certified information security or cybersecurity framework (or equivalent) is required since we demand an overall data security awareness from the tenderer. We added specific information security and data protection requirements in Annex 1 to indicate elements that would contribute to the concept of 'security-and-privacy-by-design'. These requirements do not relate to the AI Act.</p>  |

|    |         |     |   |   |   |   |
|----|---------|-----|---|---|---|---|
| 20 | Annex 5 | 4,2 | Price and other financial provisions - VAT rules      | 4 | Can we have clarification on what the VAT rules that would be applicable on this for a UK company. We understand that VAT would apply but would not be charged on our invoices. We would like to know what the Authorities view is on this in order to specify the costs                        | Answer to be disclosed in the next Q&A.   |
| 21 | Annex 5 | 4,4 | Price and other financial provisions - price schedule | 4 | We note that payment schedule for 60% of the work is upon agreement. Is there a model for this? We would like and normally, work on a milestone-based payment schedule as agreed with the customer. Is this acceptable?   | Yes this is acceptable; for the 60% portion we can define milestones and accompanying payments in the working agreements to be made by the Parties. Tenderers are invited to submit a preliminary outline (definition and planning) of subdeliverables (or milestones) as part of SAC1: Project governance plan. We have no model for this in mind. We are happy to provide a 25% advance payment. The final 15% will be paid when the assignment has been completed and all products have been accepted. |
| 22 |         |     | EU AI Act   |   | As requested in the previous market survey, we would like clarity on the different roles that apply according to the EU AI act, e.g. provider of the AI systems. Are we correct in understanding that since it is a prototype at this stage such assignments are not necessary. Please clarify. | Even though a prototype (without implementation) does not fall under the AI Act, we believe it is indeed wise to discuss the roles and responsibilities in advance. This will be a topic for discussion in the first 4 months after the contract is awarded. Tenderers are invited to submit a proposal for the division of roles and/or specification of responsibilities as part of SAC 4: Compliance plan. This part of the proposal is not binding but would serve as input for the later discussion. |