

SPEEDUP DataQuest

Réponses aux questions reçues

QUESTION 1

Should the bordereau de soumission at annexe 10.1 be signed as it stands for the present contract, or will an updated version be issued? We ask because the object it names is "marché public de services relatif à l'Assistant de recherche sur le site internet de la Chambre des Députés", and we should like to be certain we are using the intended form.

Réponse / Answer

FR : Une version mise à jour du bordereau de soumission a été rediffusée. Nous invitons les soumissionnaires à signer et à remettre cette version révisée.

EN : An updated version of the bid form has been reissued. We ask bidders to sign and submit this revised version.

QUESTION 2

Which basis should we use for sizing and pricing at section 3.3.4: the figure of approximately one million documents, or the twenty per cent of the total given in the same sentence? One million of 4,3 million corresponds to roughly 23 %, and we should like to work from the intended figure.

Réponse / Answer

FR : Pour le dimensionnement et le chiffrage de la prestation au titre de l'article 3.3.4, retenez le volume d'environ un million de documents. La référence à 20 % du volume total est donnée à titre indicatif.

EN : For the purposes of sizing and pricing under Section 3.3.4, please use the figure of approximately one million documents. The reference to 20% of the total volume is provided as an indicative proportion only.

QUESTION 3

Could you clarify what the "génération automatisée de réponses" at section 2.2 covers? We were unable to find a further definition in chapter 4 or in the acceptance criteria at section 6.4.6, and should like to confirm whether it denotes a search and question and answer interface for archivists, assisted drafting of responses to external access requests, or solely the natural language restitution into the inventory matrix described at section 3.2.4, step 7. Is it within the scope of the V1?

Réponse / Answer

FR : Le terme « génération automatisée de réponses », mentionné à la section 2.2, fait référence à la restitution en langage naturel au sein de la matrice d'inventaire décrite à la section 3.2.4, étape 7. Cette fonctionnalité s'inscrit dans le périmètre de la version 1, dans le cadre du processus de génération et d'enrichissement de l'inventaire.



EN : The term “génération automatisée de réponses” referred to in Section 2.2 relates to the natural language restitution within the inventory matrix described in Section 3.2.4, Step 7. This functionality is within the scope of V1 as part of the inventory generation and enrichment process.

QUESTION 4

Should the costs of infrastructure and hosting be included within the 140.000 EUR excluding VAT, or presented as a separate estimate alongside the price? We ask because section 9.2 requires all costs necessary for performance to be included in the lump sum price, while section 5.1.3 states that "une estimation budgétaire de l'hébergement est attendue" and selection criterion 3 asks the tenderer to present and financially evaluate the target infrastructure.

Réponse / Answer

FR : Les coûts d'infrastructure et hébergement ne sont pas inclus dans le montant maximal de 140 000 EUR hors TVA du présent marché. Ils seront pris en charge séparément par la CHD. Le soumissionnaire doit néanmoins présenter une estimation détaillée des coûts de l'infrastructure cible nécessaire au fonctionnement de sa solution. Cette estimation ne sera pas intégrée au prix de l'offre, mais sera prise en compte dans son évaluation afin d'apprécier le coût global de la solution.

EN : Infrastructure and hostings costs are not included in the maximum amount of 140,000 EUR (excluding VAT) for this contract. They will be covered separately by the CHD. The bidder must nevertheless submit a detailed estimate of the costs of the target infrastructure required for the operation of its solution. This estimate will not be included in the bid price but will be taken into account in the evaluation to assess the overall cost of the solution.

QUESTION 5

Could the CHD AI Charter be made available to tenderers? It is binding on the contractor under section 5.4.3, and we were not able to find it among the annexes listed in chapter 10.

Réponse / Answer

FR : Oui, la charte IA se trouve en libre accès sur le site internet de la CHD :

https://www.chd.lu/sites/default/files/2024-07/Charte_IA_CHD.pdf

EN : Yes, the AI Charter is freely available on the CHD website:

https://www.chd.lu/sites/default/files/2024-07/Charte_IA_CHD.pdf

QUESTION 6

What compute capacity, and in particular what GPU capacity, would the CHD or the CTIE make available, and for which of the four environments required by section 5.1.5 (development, test, acceptance and preproduction, production)? Section 5.1.3 asks tenderers to specify the "puissance de calcul attendue" and "type de coeur utilisée", and we should like to know against which baseline that sizing is to be established.



Réponse / Answer

FR : Aucune capacité de calcul ou de GPU n'est imposée ni mise à disposition à ce stade pour les différents environnements mentionnés à l'article 5.1.5. Les éléments demandés visent à permettre l'évaluation des hypothèses de dimensionnement retenues par les candidats. Ainsi, aucun dimensionnement de référence n'est prescrit par la CHD ou le CTIE.

EN : No specific compute or GPU capacity is prescribed or made available at this stage for the environments referred to in Section 5.1.5. The purpose of these requirements is precisely to allow the evaluation of the sizing assumptions proposed by tenderers. Therefore, no baseline sizing is imposed by the CHD or the CTIE.

QUESTION 7

Section 3.4.4 lists the choice of LLM service among the points still open, referring to "les solutions mises à disposition par le CTIE". Could you indicate which inference services are available today within the CTIE or the GovCloud, which models they expose, and under what access conditions?

Réponse / Answer

FR : Il n'est pas envisageable, à ce stade, de communiquer une liste des services ou modèles disponibles au sein du CTIE ou du GovCloud. La sélection définitive des LLM fait partie des points restant à préciser dans le cadre du marché.

EN : At this stage, it is not possible to provide a list of the services or models available within the CTIE or GovCloud. The final selection of LLMs is one of the points yet to be clarified under the contract.

QUESTION 8

Section 2.4.1 provides that the contractor supplies the licences necessary for operation, and section 7.3.2 asks that recurring costs be identified separately and not included in the price. For the eight months of implementation itself, are inference and model licensing costs borne by the contractor under the lump sum price, or covered by the CHD or the CTIE through an existing State framework agreement?

Réponse / Answer

FR : Les coûts de licences et d'utilisation des modèles, notamment les frais d'inférence, nécessaires à l'exécution des prestations prévues au marché doivent être inclus dans le prix forfaitaire, dans le respect du plafond de 140 000 EUR hors TVA, contingence comprise. Cela couvre l'implémentation ainsi que les opérations nécessaires à l'exécution des obligations d'hypercare. Le caractère récurrent de la facturation d'une licence, d'une souscription ou de frais d'usage ne les exclut pas du forfait lorsqu'ils sont nécessaires à l'exécution des prestations incluses dans le marché. Les coûts d'infrastructure et hébergement sont exclus de ce prix et seront pris en charge séparément par la CHD, dans le cadre d'un autre marché. Le soumissionnaire doit néanmoins fournir une estimation de ces coûts, accompagnée des hypothèses de dimensionnement retenues. Cette estimation sera prise en compte dans l'évaluation de l'architecture et du coût global de la solution, conformément à la section 8.7.3.



Les prestations d'installation, de paramétrage et d'intégration relevant du présent marché restent comprises dans le forfait. Les coûts récurrents correspondant aux prestations exclues conformément à la section 7.3.2 doivent être identifiés séparément, sans être intégrés au prix du présent marché.

EN : The costs of licences and the use of models, including inference costs, required to perform the services set out in the contract must be included in the lump sum, subject to a ceiling of EUR 140,000 excluding VAT, including the contingency. This covers implementation as well as the operations necessary to fulfil the hypercare obligations. The fact that licence, subscription or usage fees are invoiced on a recurring basis does not exclude them from the lump sum where they are necessary for the performance of the services included in the contract. Infrastructure and hosting costs are excluded from this price and will be borne separately by the CHD under a separate contract. The tenderer must nevertheless provide an estimate of these costs, accompanied by the sizing assumptions adopted. This estimate will be taken into account in the assessment of the architecture and the overall cost of the solution, in accordance with section 8.7.3. Installation, configuration and integration services covered by this contract remain included in the fixed-price package. Recurring costs relating to services excluded in accordance with section 7.3.2 must be identified separately and must not be included in the price of this contract.

QUESTION 9

Could you provide the breakdown of the approximately one million document perimeter by file type and by language, even as orders of magnitude, together with its volume in bytes? This bears directly on section 6.4.6.7, which requires that the recall gap between comparable segments not exceed three percentage points, and therefore on the feasibility of the stratified evaluation.

Réponse / Answer

FR : Le pouvoir adjudicateur n'est pas en mesure de fournir une ventilation du périmètre documentaire par type de fichier, par langue ou par volume en octets. Les soumissionnaires sont invités à formuler leurs propositions sur la base des informations figurant dans les documents de consultation. Les lots représentatifs seront définis en phase 1 avec la CHD, après caractérisation du corpus selon les formats, les langues, les volumes et les difficultés de traitement. Dans son offre, le prestataire devra proposer une méthode pour caractériser le corpus et définir les lots représentatifs, préciser les modalités de validation de la solution, notamment pour les documents difficiles et les cas limites.

EN : The contracting authority is unable to provide a breakdown of the scope of the documents by file type, language or volume in bytes. Tenderers are invited to draw up their proposals on the basis of the information contained in the tender documents. Representative batches will be defined in Phase 1 in consultation with the CHD, once the corpus has been characterised according to formats, languages, volumes and processing difficulties. In their tender, the service provider must propose a method for characterising the corpus and defining representative batches, and specify the procedures for validating the solution, particularly for difficult documents and borderline cases.



QUESTION 10

Section 4.1.4.2 refers to the quality of OCR results "lorsqu'ils sont existants", which suggests a partially processed corpus. What proportion consists of scanned documents with no usable text layer, what proportion has already been processed by OCR, what is the known quality of that existing OCR, and is running OCR again on documents already processed permitted?

Réponse / Answer

FR : Une partie du corpus est effectivement constituée de documents numérisés ayant déjà fait l'objet d'un traitement OCR. Toutefois, le pouvoir adjudicateur n'est pas en mesure, à ce stade, de fournir une estimation fiable de la répartition entre documents disposant ou non d'une couche texte exploitable, ni d'indiquer la qualité globale des traitements OCR déjà réalisés. Par ailleurs, il est autorisé de procéder à un nouveau traitement OCR sur des documents ayant déjà fait l'objet d'un traitement antérieur.

EN : Part of the corpus does indeed consist of digitised documents that have already undergone OCR processing. However, the contracting authority is not, at this stage, in a position to provide a reliable estimate of the breakdown between documents that do and do not have a usable text layer, nor to indicate the overall quality of the OCR processing already carried out. Furthermore, it is permissible to carry out a new round of OCR processing on documents that have already undergone previous processing.

QUESTION 11

What are the size and composition of the first perimeter described at section 6.5.2 as a quick win: number of files and documents, fund, format and language? We should also be grateful for a definition of the criteria listed there as "format unique, langue unique, mono-caractère, périmètre métier cohérent", the third of which we were not able to interpret with certainty.

Réponse / Answer

FR : La taille du premier périmètre décrit à la section 6.5.2 n'est pas prédéfinie par le pouvoir adjudicateur. Elle doit être déterminée par le soumissionnaire dans le cadre de l'approche qu'il propose pour atteindre les objectifs du quick win. Pour ce périmètre initial, les critères doivent être compris comme suit :

- Format unique : documents appartenant à un même format.
- Langue unique : documents rédigés dans une seule langue.
- Mono-caractère : documents utilisant un même système d'écriture ou alphabet.
- Périmètre métier cohérent : documents relevant d'un même domaine métier ou d'une même activité clairement identifiée.

EN : The scope of the first phase described in section 6.5.2 is not predefined by the contracting authority. It must be determined by the tenderer as part of the approach they propose to achieve the 'quick win' objectives. For this initial scope, the criteria are to be understood as follows:

- Single format: documents belonging to the same format.
- Single language: documents written in a single language.
- Single script: documents using the same writing system or alphabet.



- Consistent business scope: documents relating to the same business domain or a clearly identified activity.

QUESTION 12

Section 3.4.4 indicates that the useful volume is to be confirmed during the framing phase. Independently of that figure, does the contract commit the contractor to a processed volume during the eight months, or to delivery of the platform and the four increments, with processing at scale falling to subsequent operation?

Réponse / Answer

FR : Si le volume utile doit être confirmé au cours de la phase de mise en place de la structure, le marché comprend bien une phase opérationnelle de traitement des documents.

EN : Whilst the usable volume is to be confirmed during the set-up phase of the structure, the contract includes an operational phase for document processing.

QUESTION 13

Could a representative sample of the corpus be made available under the confidentiality agreement of annexe 10.2 before the submission deadline, or failing that at the start of the framing phase?

Réponse / Answer

FR : Aucun échantillon représentatif du corpus ne sera mis à la disposition des soumissionnaires avant la date limite de dépôt des offres. Le projet débutera par une phase de cadrage technique et fonctionnel visant à consolider la compréhension des besoins, à valider les choix architecturaux et à préparer les conditions de mise en œuvre de la solution. Cette phase portera notamment sur les ensembles de documents pertinents, les processus métier, l'architecture existante, les exigences de sécurité, les infrastructures et les services mis à disposition par la CHD ou le CTIE, ainsi que sur les prérequis nécessaires au déploiement de la solution.

EN : A representative sample of the corpus will not be made available to tenderers before the submission deadline. The project will begin with a technical and functional framing phase aimed at consolidating the understanding of requirements, securing architectural choices, and preparing the conditions for the implementation of the solution. This phase will cover, in particular, the relevant document collections, business processes, existing architecture, security requirements, infrastructures and services made available by the CHD or the CTIE, as well as the prerequisites necessary for deploying the solution.

QUESTION 14

Section 6.4.6.1 provides that the expected results will be "établis ou validés par la CHD". Does an annotated reference corpus exist today? Where the CHD validates rather than establishes, is the contractor expected to produce the initial annotation, over what volume, and is that effort to be included in the lump sum price?



Réponse / Answer

FR : Il n'existe pas à ce stade de corpus de référence annoté mis à disposition des soumissionnaires. Conformément à la section 6.4.6.1, chaque lot documentaire fera l'objet d'une évaluation fondée sur un corpus de référence validé par la CHD. La constitution de ce corpus et la définition des résultats attendus seront réalisées ou validées par la CHD dans le cadre du projet. L'ensemble des prestations nécessaires à la bonne exécution du marché, y compris celles liées à l'évaluation, à la préparation des jeux de données et à la mesure de la qualité des résultats, doit être pris en compte dans le prix forfaitaire.

EN : At this stage, there is no annotated reference corpus available to tenderers. In accordance with section 6.4.6.1, each data set will be assessed on the basis of a reference corpus validated by the CHD. The compilation of this corpus and the definition of the expected results will be carried out or validated by the CHD as part of the project. All services necessary for the proper performance of the contract, including those relating to evaluation, the preparation of datasets and the measurement of the quality of results, must be taken into account in the lump-sum price.

QUESTION 15

Which sensitivity categories does the CHD qualify as critical for the purposes of the recall threshold of at least 99,5 % at section 6.4.6.5?

Réponse / Answer

FR : Le CCTP ne désigne actuellement aucune des neuf catégories de données sensibles comme « critique ». Le seuil de rappel de 99,5 % s'appliquera aux catégories que la CHD qualifiera explicitement ainsi. Leur liste et les règles d'annotation devront être validées par les équipes métier et juridiques, puis intégrées au référentiel et au jeu d'évaluation de l'incrément concerné. Cette qualification ne peut pas être déduite de la seule liste des neuf catégories.

EN : The specifications do not currently designate any of the nine sensitive-data categories as "critical". The 99.5% recall threshold will apply to categories expressly designated as such by the CHD. The list and annotation rules must be validated by the business and legal teams, then included in the reference definitions and evaluation set for the relevant increment. Critical status cannot be inferred merely from inclusion in the list of nine categories.

QUESTION 16

Section 6.4.6 states that the minima apply independently to each increment and each category, while section 6.5.2 provides that sensitive data identification "pourra être limitée à des cas élémentaires et ne visera pas, à ce stade, une couverture exhaustive". How should these two provisions be read together for the first increment?

Réponse / Answer

FR : Les dispositions des sections 6.4.6 et 6.5.2 sont complémentaires. La section 6.5.2 définit un périmètre documentaire initial volontairement limité et autorise une couverture restreinte des données sensibles à des cas élémentaires. La section 6.4.6 définit quant à elle les critères de qualité applicables au périmètre retenu. Les seuils minimaux demeurent donc applicables à chaque catégorie incluse dans le premier incrément, même si ce premier incrément ne couvre



qu'un sous-ensemble limité des cas de données sensibles. Le premier incrément peut se limiter à un périmètre documentaire simple et à des cas élémentaires de données sensibles. Les seuils de la section 6.4.6 restent applicables à chaque catégorie effectivement incluse et évaluée dans cet incrément, ils ne supposent pas que les neuf catégories soient couvertes dès cette étape. Le périmètre retenu et les résultats attendus seront définis avant l'évaluation sur un corpus de référence validé par la CHD.

EN : The provisions of sections 6.4.6 and 6.5.2 are complementary. Section 6.5.2 defines an initial scope of documentation that is deliberately limited and permits restricted coverage of sensitive data to basic cases. Section 6.4.6, for its part, defines the quality criteria applicable to the selected scope. The minimum thresholds therefore remain applicable to each category included in the first increment, even if that first increment covers only a limited subset of cases involving sensitive data. The first increment may be limited to a simple document scope and elementary sensitive-data cases. The thresholds in Section 6.4.6 remain applicable to each category actually included and assessed in that increment; they do not require all nine categories to be covered at this stage. The selected scope and expected results will be defined before evaluation using a reference corpus validated by the CHD.

QUESTION 17

Could the "document de référence des définitions" of the sensitive cases, announced at section 3.1.5.2, be made available to tenderers, or failing that, by what date will it be available?

Réponse / Answer

FR : Le document sera disponible après la contractualisation avec le titulaire (pour la version provisoire, la version finale est prévue pour février 2027). Le tableau présent dans le CCTP est le livrable de référence jusqu'alors.

EN : The document will be available once the contract with the contractor has been signed (for the draft version; the final version is scheduled for February 2027). Until then, the table included in the Technical Specifications serves as the reference document.

QUESTION 18

Section 5.4.3 provides that the contractor contributes to the regulatory qualification of the solution. Has the CHD already formed a preliminary view as to whether the system falls within the high risk category of the AI Act, and how are the roles of provider and deployer allocated between the CHD and the contractor? Given section 9.6, under which the contract cannot be modified or adapted, how would additional obligations arising from a qualification established after signature be handled?

Réponse / Answer

FR : L'analyse de risques effectuée par la CHD démontre que le cas d'usage n'est pas à « à haut risque ». La CHD utilisera la solution sous son autorité, la qualité de fournisseur dépendra notamment de l'entité qui mettra le système en service sous son nom. L'installation par le titulaire ne fait pas de lui le déployeur. La section 9.6 ne prévoit aucun ajustement automatique



du prix ou du calendrier : toute conséquence nouvelle devra être examinée avant déploiement dans le cadre juridique du marché.

EN : The risk analysis conducted by the CHD shows that the use case is not “high risk.” The CHD will use the solution under its own authority; the quality of the supplier will depend, in particular, on the entity that commissions the system under its own name. Installation by the license holder does not make it the deployer. Section 9.6 does not provide for any automatic adjustment to the price or schedule: any new consequences must be reviewed prior to deployment within the legal framework of the contract.

QUESTION 19

Does the prohibition at section 9.7 on "services de cloud privés" cover any external SaaS or API service, including a Mistral service hosted within the European Union, or only clouds outside the government sector? We ask in view of section 3.2.3, which presents Mistral technologies as a reference target.

Réponse / Answer

FR : L'hébergement de la solution sera réalisé par la CHD. Les caractéristiques et les partenaires seront définis durant le cadrage. Un modèle Mistral peut être proposé dans une architecture conforme aux exigences d'hébergement et de sécurité de la CHD. Le soumissionnaire devra décrire le service envisagé et les flux de données associés.

EN : The solution will be hosted by the CHD. The specifications and partners will be defined during the scoping phase. A Mistral model may be proposed within an architecture that complies with the CHD's hosting and security requirements. The bidder must describe the proposed service and the associated data flows.

QUESTION 20

Section 8.7.2 asks that effort be quantified "en termes de charge de travail, de profils mobilisés". Should the interchangeability effort required by selection criterion 2 be expressed in the same unit, in cost, or in both? Which components does the CHD consider critical in this respect: LLM, OCR engine, orchestrator, storage and index, restitution layer?

Réponse / Answer

FR : L'exigence relative à l'interchangeabilité des composants de l'architecture relève du critère de sélection concerné (section 8.6) et est distincte du critère d'attribution n° 2 (visé à la section 8.7.2). La quantification de l'effort d'interchangeabilité est attendue sous une forme suffisamment détaillée pour permettre d'en évaluer l'ampleur et la faisabilité. Le coût correspondant est réputé inclus dans l'offre financière du soumissionnaire. Aucun composant n'est considéré comme critique de manière prédéfinie. Les soumissionnaires sont invités à identifier les composants qu'ils jugent pertinents du point de vue de l'interchangeabilité.

EN : The requirement regarding the interchangeability of architectural components falls under the relevant selection criterion (section 8.6) and is distinct from award criterion No. 2 (referred to in section 8.7.2). The effort required to achieve interchangeability is to be quantified in



sufficient detail to enable an assessment of its scope and feasibility. The corresponding cost is deemed to be included in the tenderer's financial offer. No component is considered critical by default. Tenderers are invited to identify the components they consider relevant from an interchangeability perspective.

QUESTION 21

The data protection agreement at annexe 10.4 gives the CHD three months from notification to object to the addition or replacement of a subprocessor. Services start on 11 January 2027 and section 6.5.2 targets a usable first perimeter by mid February 2027, while section 3.4.4 leaves the LLM service and the hosting arrangements to be settled during the framing phase. How is this objection period intended to operate for subprocessors identified during framing, and may providers named in the tender be approved at contract signature?

Réponse / Answer

FR : Le même délai de trois mois s'applique également pour la période « framing », tous les sous-traitants nommés dans l'offre seront normalement acceptés lors de la signature du contrat (sous réserve des conditions imposées aux sous-traitants par le cahier de charges). Pour le reste, merci de se référer à la réponse faite aux questions 8, 24 et 72.

EN : The same three-month deadline also applies to the 'framing' period; all subcontractors named in the tender will normally be accepted upon signature of the contract (subject to the conditions imposed on subcontractors by the tender specifications). For the rest, please refer to the answers to questions 8, 24 and 72.

QUESTION 22

The same annexe requires personal data breaches to be notified within two hours of the contractor becoming aware of them, with no qualification as to working hours, whereas section 7.2.1 defines hypercare availability within the CHD's business hours. Should the two hour obligation be understood as applying on a continuous basis?

Réponse / Answer

FR : L'obligation de notification des violations de données est distincte du dispositif d'hypercare. La notification d'une violation en relation avec des données personnelles doit être notifiée endéans un délai de 2h après prise de connaissance, nonobstant si cette prise de connaissance se fait hors heures d'ouvertures « normales ».

EN : The obligation to report data breaches is separate from the hypercare scheme. A breach involving personal data must be reported within two hours of it coming to light, regardless of whether this occurs outside 'normal' working hours.

QUESTION 23

Section 1.14 requires all members of a collective offer to satisfy the selection criteria. Should this be read as each member individually evidencing three references, or as the grouping evidencing three collectively? Where subcontracting is used instead, may references held by a declared



subcontractor be taken into account? We note in addition that annexe 10.2, article 4, requires the contractor not to engage any subcontractor without the prior written agreement of the Chamber, and should be grateful for confirmation that subcontractors declared in the tender under section 1.13 are deemed approved for that purpose.

Réponse / Answer

FR : Oui. Pour un groupement, les références peuvent être présentées collectivement par ses membres, avec un minimum de trois références au total. Les références d'un sous-traitant déclaré dans l'offre peuvent également être prises en compte. Conformément à la section 8.6 : « Le candidat devra justifier d'au moins trois références de projets similaires. Les projets présentés décriront le même type de technologie. Ils porteront soit sur l'analyse de corpus documentaires hétérogènes, soit sur des organismes publics. » Par ailleurs, les sous-traitants dûment déclarés dans l'offre conformément à la section 1.13 et acceptés dans le cadre de l'attribution du marché seront réputés avoir reçu l'accord écrit préalable de la Chambre au sens de l'annexe 10.2, article 4.

EN : Yes. In the case of a consortium, the references may be submitted collectively by its members, with a minimum of three references in total. References held by a subcontractor declared in the tender may also be taken into account. In accordance with Section 8.6: "The candidate must provide evidence of at least three references for similar projects. The projects presented must use the same type of technology. They must concern either the analysis of heterogeneous documentary corpora or public-sector bodies." Furthermore, subcontractors duly declared in the tender pursuant to Section 1.13 and accepted as part of the contract award will be deemed to have received the Chamber's prior written approval for the purposes of Annex 10.2, Article 4.

QUESTION 24

Section 8.7.2 requires the tenderer to establish the link between the requirements of the Cahier des charges and the functionalities proposed. Given the eight A4 page limit applying to parts 1 to 8, may a requirements traceability matrix be submitted as a separate annex outside that limit?

Réponse / Answer

FR : Oui. Les soumissionnaires peuvent joindre une matrice de traçabilité facultative en annexe du mémoire principal. Cette annexe n'est pas comptabilisée dans la limite des huit pages prévue à la section 8.9. La matrice permet de présenter de manière synthétique la correspondance entre les exigences du CCTP et les fonctionnalités ou prestations proposées. Elle indique la référence de chaque exigence, la couverture proposée et le renvoi précis au passage correspondant du mémoire principal. Elle distingue les éléments couverts de manière standard, ceux nécessitant une adaptation et les éventuelles couvertures partielles ou exclusions. Des exigences faisant l'objet d'une même réponse peuvent être regroupées, à condition que leurs références restent clairement identifiables. Le mémoire principal doit contenir, dans la limite des huit pages, les explications de fond sur la solution et l'approche proposées, ainsi que leurs justifications, hypothèses, limites et exclusions. La matrice complète la présentation de la couverture des exigences ; elle ne peut servir à ajouter des développements techniques ou méthodologiques absents du mémoire principal. Cette annexe est limitée à cinq pages A4 au total, en orientation



portrait ou paysage, au format PDF, avec une taille de caractères d'au moins 10 points. Une version tableur contenant strictement les mêmes informations peut également être jointe. Le PDF constitue la version de référence de cette annexe. L'obligation de justifier la couverture des exigences demeure applicable avec ou sans cette matrice. Cette autorisation porte uniquement sur la matrice de traçabilité.

EN : Yes. Tenderers may attach an optional traceability matrix as an annex to the main tender document. This annex is not counted towards the eight-page limit set out in section 8.9. The matrix provides a summary of how the requirements of the tender specifications correspond to the proposed functionalities or services. It sets out the reference for each requirement, the proposed coverage and a precise cross-reference to the relevant section of the main proposal. It distinguishes between elements covered as standard, those requiring adaptation, and any partial coverage or exclusions. Requirements to which the same response applies may be grouped together, provided that their references remain clearly identifiable. The main proposal must contain, within a limit of eight pages, detailed explanations of the proposed solution and approach, together with their justifications, assumptions, limitations and exclusions. The matrix supplements the presentation of the coverage of the requirements; it must not be used to add technical or methodological details not included in the main proposal. This annexe is limited to a total of five A4 pages, in portrait or landscape orientation, in PDF format, with a font size of at least 10 points. A spreadsheet version containing exactly the same information may also be attached. The PDF constitutes the reference version of this annexe. The obligation to demonstrate compliance with the requirements remains in force regardless of whether this matrix is used or not. This authorisation relates solely to the traceability matrix.

QUESTION 25

Criteria 1, 2, 3 and 5 total 85 points and criterion 4 carries 15, giving 100. Should the five points attached to the optional demonstration at section 8.7.4 be understood as included within the fifteen points of criterion 4, such that a candidate not giving a demonstration may obtain at most ten, or as genuinely additional?

Réponse / Answer

FR : Les cinq points associés à la démonstration facultative prévue à la section 8.7.4 sont des points bonus, qui s'ajoutent aux 100 points attribués au titre des cinq critères.

EN : The five points associated with the optional demonstration provided for in Section 8.7.4 are bonus points, awarded in addition to the 100 points available under the five criteria.

QUESTION 26

Le cahier mentionne environ 4,3 millions de documents et estime à environ 1 million le volume potentiellement soumis aux traitements IA. Pouvez-vous confirmer l'ordre de grandeur attendu et préciser si l'unité considérée est le fichier, la pièce documentaire ou le dossier ?

Réponse / Answer



FR : Pour la volumétrie, au titre de l'article 3.3.4, reprenez le volume d'environ un million de documents potentiellement soumis au traitement IA dans le cadre du projet. L'analyse pourra être menée au niveau des pièces, mais devra notamment produire une restitution consolidée au niveau du dossier. Le marché porte toutefois sur la mise en place d'une solution opérationnelle, et non sur le traitement exhaustif de ce corpus pendant le projet. Le titulaire devra donc déterminer et justifier le volume de documents nécessaire à la conception, à l'entraînement et à la validation de la solution, afin d'atteindre les niveaux de qualité attendus et de limiter les erreurs. Ce volume pourra porter sur tout ou partie du corpus, selon les besoins démontrés. À l'issue du projet, la CHD devra pouvoir traiter le solde des documents de manière autonome au moyen de la solution mise en place.

EN : With regard to volume, in accordance with Article 3.3.4, assume a volume of approximately one million documents that may be subject to AI processing as part of the project. The analysis may be carried out at document level, but must, in particular, produce a consolidated output at file level. However, the contract relates to the implementation of an operational solution, and not to the exhaustive processing of this corpus during the project. The contractor must therefore determine and justify the volume of documents required for the design, training and validation of the solution, in order to achieve the expected quality levels and minimise errors. This volume may cover all or part of the corpus, depending on the demonstrated requirements. Upon completion of the project, the CHD must be able to process the remaining documents autonomously using the solution put in place.

QUESTION 27

Pour l'objectif de 100 000 documents traités en maximum 8 heures, pouvez-vous préciser si le temps mesuré comprend le contrôle de recevabilité, l'extraction, l'OCR éventuel, les appels IA, les reprises et la génération des résultats ? Le corpus et l'infrastructure de référence seront-ils fournis ou définis pendant le cadrage ?

Réponse / Answer

FR : Les huit heures couvrent la chaîne automatisée depuis l'acceptation du lot jusqu'à la disponibilité des résultats, rapports et statuts. La mesure comprend donc l'extraction, l'OCR lorsqu'il est nécessaire, les traitements IA, les contrôles automatisés et les reprises techniques. Elle exclut les délais de validation humaine. Le corpus de performance et l'infrastructure de référence seront validés lors du cadrage et pourraient évoluer. Le titulaire devra justifier ses hypothèses et mesurer le temps total ainsi que le temps par étape.

EN : The eight-hour period covers the automated chain from batch acceptance until results, reports and statuses are available. It therefore includes extraction, OCR where needed, AI processing, automated checks and technical retries; human validation time is excluded. The performance corpus and reference infrastructure will be validated during scoping. The contractor must justify its assumptions and measure both the total duration and the time spent at each stage.



QUESTION 28

Les 19 formats mentionnés doivent-ils tous bénéficier de l'intégralité des traitements dès la V1, ou leur prise en charge peut-elle être étendue progressivement au cours des quatre incréments ?

Réponse / Answer

FR : La solution doit permettre la prise en charge de l'ensemble des formats dès la V1. La découpe par format n'est pas possible car elle ne correspond pas à la réalité des fonds qui sont hétérogènes. Des lots représentatifs seront définis par le prestataire avec la CHD pour chaque incrémentation.

EN : The solution must support all formats from version 1 onwards. It is not possible to categorise by format, as this does not reflect the reality of the collections, which are heterogeneous. Representative batches will be defined by the service provider in consultation with the CHD for each increment.

QUESTION 29

Serait-il possible de disposer avant la remise de l'offre d'un exemple anonymisé ou d'un modèle représentatif de la matrice d'inventaire afin d'en connaître précisément les feuilles, colonnes, types de données et contraintes ?

Réponse / Answer

FR : La matrice d'inventaire est décrite précisément dans le CCTP. Elle ne pourra pas être transmise avant la remise de l'offre.

EN : The inventory matrix is described in detail in the Specifications. It cannot be provided until the tender has been submitted.

QUESTION 30

Le document de référence complet contenant les définitions des catégories de données sensibles sera-t-il communiqué avant la remise de l'offre ou uniquement pendant la phase de cadrage ?

Réponse / Answer

FR : Le document sera disponible après la contractualisation avec le titulaire. Le tableau présent dans le CCTP est le livrable de référence jusqu'alors.

EN : The document will be available once the contract with the contractor has been finalised. Until then, the table contained in the Specifications is the reference document.

QUESTION 31

Disposez-vous déjà d'indications concernant l'environnement privilégié (CTIE, CHD/GovCloud ou autre environnement souverain autorisé), ainsi que les modalités envisagées pour l'accès à Mistral ?



Réponse / Answer

FR : L'ensemble des indications relatives aux environnements sont décrites dans le CCTP. Il appartient au titulaire de proposer et de mettre en œuvre la solution technique répondant aux exigences du marché, y compris les modalités d'accès aux services d'IA tels que Mistral, dans le respect des contraintes et exigences définies dans les documents du marché

EN : All specifications relating to the environments are set out in the Technical Specifications. It is the contractor's responsibility to propose and implement the technical solution that meets the tender requirements, including the arrangements for accessing AI services such as Mistral, in accordance with the constraints and requirements set out in the tender documents.

QUESTION 32

Pouvez-vous confirmer que DataQuest doit se concentrer sur l'analyse, l'enrichissement, la qualification et l'assistance aux archivistes, sans reprendre les fonctions de gestion documentaire, de versement et de conservation relevant de COMED ? Une fonction de recherche et de consultation des documents analysés est-elle souhaitée dans DataQuest ou doit-elle rester hors périmètre de la V1 ?

Réponse / Answer

FR : Le périmètre de DataQuest est précisé au chapitre 3.1.3. Le projet COMED est présenté à titre de contexte et ne fait pas partie du périmètre du présent marché. La fonction de recherche et consultation n'est pas souhaitée.

EN : The scope of DataQuest is set out in section 3.1.3. The COMED project is presented for background purposes only and does not form part of the scope of this contract. The search and browse function is not required.

QUESTION 33

Comment faut-il articuler l'exigence de suppression intégrale des copies après production de l'inventaire avec la conservation nécessaire des preuves, résultats, journaux, sauvegardes et éléments permettant la reproductibilité et l'audit pendant le BUILD et l'hypercure ?

Réponse / Answer

FR : L'exigence de suppression des copies vise les documents et jeux de données constitués pour les besoins du traitement. Elle ne fait pas obstacle à la conservation, pendant les phases de BUILD et d'hypercure, des éléments strictement nécessaires à l'exécution du contrat, à la traçabilité des traitements, à l'auditabilité de la solution et à la reproductibilité des résultats. Ces éléments devront être limités au nécessaire et gérés conformément aux exigences de sécurité, de confidentialité et de protection des données prévues par le marché.

EN : The requirement to delete copies applies to documents and datasets created for the purposes of processing. It does not preclude the retention, during the BUILD and hypercare phases, of items strictly necessary for the performance of the contract, the traceability of processing operations, the auditability of the solution and the reproducibility of results. These



elements must be limited to what is strictly necessary and managed in accordance with the security, confidentiality and data protection requirements set out in the contract.

QUESTION 34

Could the CHD provide additional information on the characteristics and distribution of the documentary corpus:

- number of documents/files by type and format;
- typical and maximum document size and/or number of pages;
- proportion of scanned documents requiring OCR versus digitally generated documents;
- degree of standardisation or use of recurring document templates;
- representative volumes expected to be processed within each increment?

Réponse / Answer

FR : Le pouvoir adjudicateur ne dispose pas, à ce stade, d'informations suffisamment consolidées permettant de fournir une répartition fiable du corpus selon les types de documents, formats, tailles, volumes, taux de numérisation/OCR ou niveaux de standardisation. Les soumissionnaires sont invités à construire leur offre sur la base des informations quantitatives et qualitatives disponibles dans les documents du marché.

EN : At this stage, the contracting authority does not have sufficiently comprehensive information to provide a reliable breakdown of the corpus by document type, format, size, volume, digitisation/OCR rate or level of standardisation. Tenderers are invited to prepare their tenders on the basis of the quantitative and qualitative information available in the tender documents.

QUESTION 35

Regarding the 100K documents within 8 hours. How should this requirement be interpreted for sizing purposes? Will the contractual performance test be based exclusively on the reference corpus and infrastructure agreed during scoping?

Réponse / Answer

FR : L'exigence de traitement de 100 000 documents en 8 heures constitue un objectif de performance à atteindre dans le cadre de la solution proposée. Les soumissionnaires sont invités à dimensionner leur offre en conséquence, en tenant compte des hypothèses et contraintes décrites dans le cahier des charges. Les soumissionnaires sont invités à présenter la démonstration d'un outil jugé pertinent pour démontrer l'atteinte d'un objectif similaire à celui de la CHD.

EN : The requirement to process 100,000 documents within 8 hours constitutes a performance target to be met as part of the proposed solution. Tenderers are invited to tailor their proposals accordingly, taking into account the assumptions and constraints set out in the tender specifications. Tenderers are invited to provide a demonstration of a tool deemed appropriate to demonstrate the achievement of a target similar to that of the CHD.



QUESTION 36

The RfO specifies an implementation period of eight months and four successive functional increments of approximately 60 days each. It also indicates that scoping, infrastructure preparation, platform installation and configuration of the first processing workflows will be performed in parallel.

- How the approximately 60-day duration of each increment should be interpreted within the overall eight-month BUILD period?
- In particular, is the expectation that scoping and platform setup overlap primarily with the first increment, or should the duration of each phase/increment be shortened to accommodate the initial scoping and platform setup activities?

Réponse / Answer

FR : Le calendrier présenté est indicatif. Les soumissionnaires sont libres de proposer l'organisation et le phasage qu'ils jugent les plus adaptés, dans le respect des objectifs et du délai global du projet. Une parallélisation des activités est possible et les soumissionnaires sont invités à apporter leur valeur de conseil dans la définition du planning de mise en œuvre.

EN : The timetable set out is indicative. Tenderers are free to propose the organisation and phasing they consider most appropriate, provided they comply with the project's objectives and overall deadline. Activities may be carried out in parallel, and tenderers are invited to contribute their expertise in drawing up the implementation schedule.

QUESTION 37

Interaction between hypercare and subsequent increments:

- How are issues identified during hypercare expected to interact with the ongoing BUILD activities? For example, should a non-conformity identified during the hypercare of Increment N normally be corrected immediately within that deployed perimeter, or may its correction be incorporated into the development/release of Increment N+1 where appropriate?
- In the event of a significant or blocking issue, is the subsequent increment expected to continue according to the planned schedule, or may its development/deployment be postponed until the issue has been resolved?
- How this four-week validation period is expected to operate within the three-month hypercare period, particularly when corrective actions occur late in the hypercare period?

Réponse / Answer

FR : Les activités de BUILD et d'hypercare peuvent se dérouler en parallèle. Toute non-conformité identifiée pendant l'hypercare devra être analysée et traitée selon sa criticité, conformément aux exigences définies à la section 7.2. Les corrections nécessaires à la mise en conformité du périmètre déployé relèvent de l'hypercare et ne peuvent pas être reportées systématiquement à un incrément ultérieur. Les délais seront mesurés dans l'outil de suivi depuis le signalement ou la détection automatique jusqu'à la prise en charge, au rétablissement et à la résolution. Toute suspension de délai devra être justifiée et acceptée par la CHD. La période de quatre semaines mentionnée pour la clôture de l'hypercare correspond à une période d'observation durant



laquelle les critères de qualité doivent être respectés. La clôture suppose l'absence d'incident critique ouvert, la résolution ou l'acceptation formelle des autres anomalies, le respect des seuils de qualité sur une période de quatre semaines, la réussite des tests de non-régression et la capacité de la CHD à administrer le périmètre.

EN : BUILD and hypercare activities may take place in parallel. Any non-conformity identified during hypercare must be analysed and addressed according to its criticality, in accordance with the requirements set out in section 7.2. The corrections required to bring the deployed scope into compliance fall within the remit of hypercare and cannot be systematically deferred to a subsequent increment. Timeframes will be measured in the tracking tool from the time of reporting or automatic detection through to handling, restoration and resolution. Any extension of a timeframe must be justified and approved by the CHD. The four-week period mentioned for the closure of hypercare corresponds to an observation period during which the quality criteria must be met. Closure requires that there are no open critical incidents, that other anomalies have been resolved or formally accepted, that quality thresholds have been met over a four-week period, that non-regression tests have been passed, and that the CHD is able to manage the scope.

QUESTION 38

The RfO indicates that the classification plan is currently partial and undergoing redesign, while the inventory matrix may also evolve. These elements, however, constitute important inputs and, in the case of the classification plan, a prerequisite for the automated processing.

- What would be the expected maturity and availability of these elements at the start of the project and for each increment?
- Could CHD confirm that the contractor's responsibility is to integrate, configure and operationalise the classification plan and inventory matrix provided by CHD, and not to define the archival classification framework or complete the inventory structure as a business deliverable?
- If the classification plan, inventory matrix, sensitive-data definitions or business rules are incomplete or change during BUILD, will the resulting impact on planning, effort and acceptance be handled through the evolution/arbitration process?

Réponse / Answer

FR : C'est à la CHD qu'appartient la charge de produire et mettre à disposition un plan de classement et une matrice d'inventaire pour le titulaire. Le plan de classement et la matrice d'inventaire seront communiqués au titulaire au début de chaque itération. Ainsi, la définition du plan de classement et de la matrice d'inventaire relève de la CHD. Le titulaire devra intégrer et paramétrer les versions fournies pour les traitements prévus, sans avoir à produire le référentiel archivistique comme livrable métier. Le plan étant en refonte et la matrice susceptible d'évoluer, le cadrage devra identifier les versions disponibles et les prérequis de chaque incrément. Une modification ultérieure ayant un effet sur la charge, les délais ou la recette fera l'objet d'une analyse d'impact et d'un arbitrage, sans modification automatique du marché.



EN : It is the responsibility of the CHD to produce and provide a classification plan and an inventory matrix for the contractor. The classification plan and inventory matrix will be provided to the contractor at the start of each iteration. The CHD is responsible for defining the classification plan and inventory matrix. The contractor must integrate and configure the versions supplied for the planned processing, without producing the archival framework as a business deliverable. As the plan is being redesigned and the matrix may evolve, scoping must identify the available versions and the prerequisites for each increment. A subsequent change affecting effort, timing or acceptance will be assessed and arbitrated, without automatically modifying the contract.

QUESTION 39

Scope Control:

- The RfO describes an extensive set of functional features and capabilities to be delivered within the eight-month BUILD period. Could the CHD provide the definitive list of features that are in scope and out of scope for the BUILD phase?
- Once an increment has been specified, configured, tested and accepted, how will subsequent requests to modify rules, prompts, thresholds, workflows, classification criteria, inventory fields or output formats be treated? Will these be considered corrective activities, normal configuration activities, or changes to the agreed scope?
- At what point is the functional perimeter of an increment considered frozen for acceptance purposes?

Réponse / Answer

FR : La liste des fonctionnalités incluses dans le périmètre de la phase BUILD est décrite dans le CCTP. Les incréments sont cumulatifs : les modifications apportées sont des évolutions censées se cumuler et non modifier les règles précédentes, comme évoqué dans le 6.5.1 – Axes de complexification du mémoire technique. Chaque nouveau lot devra s'appuyer sur les paramétrages, retours d'expérience et indicateurs de qualité issus des lots précédents. Le périmètre fonctionnel d'un incrément est considéré comme figé aux fins d'acceptation selon les critères énoncés en 6.4.6.13 – Décision d'admission.

EN : The list of features included within the scope of the BUILD phase is set out in the RfO. The increments are cumulative: the changes made are intended to build upon one another rather than alter previous rules, as outlined in section 6.5.1 – Axes de complexification in the technical specification. Each new lot must build upon the parameters, lessons learnt and quality indicators derived from previous lots. The functional scope of an increment is considered fixed for the purposes of acceptance in accordance with the criteria set out in 6.4.6.13 – Décision d'admission.

QUESTION 40

Infrastructure responsibility and available CTIE infrastructure :

- What infrastructure will be made available by the CHD/CTIE for the project (compute, GPU, storage, networking, backup, monitoring, etc.)?



- Is the infrastructure required to execute AI models expected to be included in the tenderer's financial proposal, or should the tenderer only specify the required infrastructure for provisioning by the CHD/CTIE?
- For availability and recovery objectives, how will interruptions attributable to CHD/CTIE infrastructure, networks, external AI/model services or other dependencies outside the contractor's control be treated?

Réponse / Answer

FR : La CHD mettra en place l'infrastructure selon les estimations du prestataire. Le prestataire doit donc préciser les coûts d'infrastructure qui sont pris en compte dans l'évaluation même s'ils feront l'objet d'un marché à part. Le scénario d'hébergement et les moyens éventuellement mis à disposition seront arrêtés lors de la phase de cadrage technique. Le soumissionnaire devra donc préciser, pour chaque scénario, les prérequis d'infrastructure nécessaires (calcul, GPU, mémoire, stockage/IOPS, réseau, sauvegarde, supervision, sécurité, etc.), distinguer clairement les coûts de mise en œuvre des coûts d'hébergement et d'infrastructure, fournir une estimation budgétaire pour les scénarios dans lesquels l'infrastructure ou les services d'IA doivent être fournis par le titulaire ou par un opérateur souverain. L'offre financière devra ainsi permettre au pouvoir adjudicateur d'évaluer et de comparer objectivement les différents scénarios, sans préjuger du choix qui sera effectué à l'issue du cadrage.

EN : The CHD will set up the infrastructure in accordance with the service provider's estimates. The service provider must therefore specify the infrastructure costs that are taken into account in the evaluation, even if they will be the subject of a separate contract. The hosting scenario and any resources that may be made available will be finalised during the technical scoping phase. The tenderer must therefore specify, for each scenario, the necessary infrastructure requirements (computing power, GPUs, memory, storage/IOPS, network, backup, monitoring, security, etc.), clearly distinguish between implementation costs and hosting and infrastructure costs, and provide a budget estimate for scenarios in which the infrastructure or AI services are to be provided by the successful tenderer or by a sovereign operator. The financial proposal must therefore enable the contracting authority to objectively assess and compare the various scenarios, without prejudging the choice that will be made following the scoping exercise.

QUESTION 41

Data preparation responsibility:

- Where is the responsibility boundary between preparation/remediation of the documentary corpus by the CHD and preprocessing performed by the solution? For example, who is responsible for correcting corrupted files, restructuring inconsistent directories, resolving missing documents or metadata, and remediating poor-quality scans?
- Will the CHD provide validated reference datasets / ground truth for evaluating Thematic DataQuest and Sensitive DataQuest, or is the contractor expected to construct these datasets together with the archivists?

Réponse / Answer



FR : Conformément à l'article 4.3.2 du CCTP, la CHD met à disposition le corpus documentaire, le plan de classement et la matrice. En collaboration avec les archivistes de la CHD, le titulaire constitue les ensembles de données et met en œuvre, si nécessaire, les traitements préparatoires convenus : nettoyage, normalisation, restructuration ou exclusion. Lors du cadrage, les parties établissent l'état initial du corpus, les seuils de qualité et les exclusions. L'objectif prévu par le CCTP est de disposer d'un outil capable de produire l'inventaire avec le minimum d'erreurs, et non de procéder à la correction manuelle de ses livrables. Toute erreur constatée devra conduire à l'ajustement de la solution ou de ses paramètres, afin que son fonctionnement ne dépende pas de corrections manuelles.

EN : In accordance with clause 4.3.2 of the Technical Specifications, the CHD shall provide the document collection, the classification scheme and the matrix. In collaboration with the CHD's archivists, the contractor shall compile the datasets and, where necessary, carry out the agreed preparatory processing: cleaning, standardisation, restructuring or exclusion. During the scoping phase, the parties establish the initial state of the corpus, the quality thresholds and the exclusions. The objective set out in the Technical Specifications is to have a tool capable of producing the inventory with the minimum number of errors, rather than having to manually correct its outputs. Any errors identified must lead to adjustments to the solution or its parameters, so that its operation does not depend on manual corrections.

QUESTION 42

Acceptance criteria for AI quality:

- Which quantitative quality criteria will determine acceptance of the AI components for each use case? In particular, will minimum precision, recall, F1 or other thresholds be defined by the CHD before implementation, or jointly established during scoping?
- Will acceptance criteria differ by sensitive-data category, document type, language or corpus?

Réponse / Answer

FR : Le titulaire devra mettre en œuvre des composants d'IA capables de respecter les seuils de qualité, et la volumétrie exprimée dans le CCTP. Les seuils de confiance sont exprimés au chapitre 4.3.3 : ils seront précisés avec le titulaire lors de la phase de cadrage. La solution devra permettre d'adapter ces seuils au fil du temps en fonction des retours d'expérience et des exigences métier.

EN : The contractor must implement AI components capable of meeting the quality thresholds and volume requirements set out in the Specifications. The confidence thresholds are set out in section 4.3.3: these will be finalised with the contractor during the scoping phase. The solution must allow these thresholds to be adjusted over time in line with feedback and business requirements.

QUESTION 43

Volumétrie (§3.3.4). Le CCTP estime à environ 1 million de documents (20 % du total de 4,3 millions) le périmètre de travail pour le traitement par IA, à confirmer lors du cadrage. Un chiffre plus précis est-il disponible dès à présent ? Le corpus pilote du premier incrément sera-t-il



sélectionné avant la signature du contrat et, le cas échéant, quel est son ordre de grandeur attendu ?

Réponse / Answer

FR : Au titre de l'article 3.3.4 du CCTP, le périmètre de référence est d'environ un million de documents potentiellement soumis au traitement par la solution d'IA. Aucune volumétrie ni précision supplémentaire n'est fixée pour le premier incrément. Le marché porte toutefois sur la mise en place d'une solution opérationnelle, et non sur le traitement exhaustif de ce corpus pendant le projet. Le titulaire devra donc déterminer et justifier le volume de documents nécessaire à la conception, à l'entraînement et à la validation de la solution, afin d'atteindre les niveaux de qualité attendus et de limiter les erreurs. Ce volume pourra porter sur tout ou partie du corpus, selon les besoins démontrés. À l'issue du projet, la CHD devra pouvoir traiter le solde des documents de manière autonome au moyen de la solution mise en place.

EN : Under Article 3.3.4 of the Technical Specifications, the reference scope comprises approximately one million documents that may be processed by the AI solution. No specific volume or further details are specified for the first phase. However, the contract relates to the implementation of an operational solution, and not to the exhaustive processing of this corpus during the project. The contractor must therefore determine and justify the volume of documents required for the design, training and validation of the solution, in order to achieve the expected quality levels and minimise errors. This volume may cover all or part of the corpus, depending on the demonstrated requirements. Upon completion of the project, the CHD must be able to process the remaining documents autonomously using the solution put in place.

QUESTION 44

Taxonomie des données sensibles (§3.1.5.2). Les définitions des neuf catégories de données sensibles sont présentées comme provisoires, en attente de validation juridique/DPO. Les soumissionnaires doivent-ils chiffrer sur la base de la version figurant dans le CCTP, ou une version plus définitive est-elle déjà en préparation ? La CHD pourrait-elle partager quelques exemples anonymisés par catégorie, afin de calibrer l'approche de détection ? Enfin, si la taxonomie évolue de manière significative après la signature du contrat, ce changement serait-il traité comme relevant du mécanisme de contingence (§1.18), ou est-il prévu qu'il soit absorbé dans le prix forfaitaire ?

Réponse / Answer

FR : Les soumissionnaires doivent chiffrer sur la base de la version figurant dans le CCTP. La CHD ne partage pas d'exemple anonymisé. Les évolutions devront être absorbées dans le prix forfaitaire. Le document sera disponible après la contractualisation avec le titulaire (pour la version provisoire, la version finale est prévue pour février 2027).

EN : Tenderers must prepare their quotations on the basis of the version set out in the Technical Specifications. The CHD does not provide anonymised examples. Any changes must be factored into the lump-sum price.



QUESTION 45

Hébergement et routage du service IA (§5.1.3, §3.2.3). Parmi les trois scénarios d'hébergement décrits (CTIE, CHD, opérateur souverain), une direction est-elle déjà privilégiée ? Par ailleurs, concernant le service Mistral AI mentionné au §3.2.3, DataQuest y accéderait-il via une infrastructure gérée par le CTIE ou via un arrangement direct ? Et puisque les offres doivent présenter et chiffrer l'infrastructure cible (§8.6, critère 3), le catalogue de services et la tarification du CTIE sont-ils accessibles aux soumissionnaires avant le dépôt des offres ?

Réponse / Answer

FR : Aucune direction n'est privilégiée parmi les 3 scénarios d'hébergement. Les technologies proposées par Mistral AI constituent une orientation privilégiée au sein de l'écosystème numérique de l'État luxembourgeois et pourront être considérées comme une cible de référence pour l'implémentation des cas d'usage prévus dans le cadre du présent projet. Aucune information supplémentaire n'est communiquée par la CHD, ni le CTIE. Les coûts d'infrastructure ne sont pas inclus dans le montant maximal de 140 000 EUR hors TVA du présent marché. Ils seront pris en charge séparément par la CHD. Le soumissionnaire doit néanmoins présenter une estimation détaillée des coûts de l'infrastructure cible nécessaire au fonctionnement de sa solution. Cette estimation ne sera pas intégrée au prix de l'offre, mais sera prise en compte dans son évaluation afin d'apprécier le coût global de la solution.

EN : No particular approach is favoured amongst the three hosting scenarios. The technologies proposed by Mistral AI represent a preferred approach within the Luxembourg government's digital ecosystem and may be regarded as a benchmark for the implementation of the use cases envisaged under this project. No further information has been provided by either the CHD or the CTIE. Infrastructure costs are not included in the maximum amount of 140,000 EUR (excluding VAT) for this contract. They will be covered separately by the CHD. The bidder must nevertheless submit a detailed estimate of the costs of the target infrastructure required for the operation of its solution. This estimate will not be included in the bid price but will be taken into account in the evaluation to assess the overall cost of the solution.

QUESTION 46

Dépendances amont (§3.2.1.1, §3.1.6). Deux éléments en amont de DataQuest sont encore en évolution. D'une part, le plan de classification de la CHD est décrit comme partiel et en cours de refonte active : une version stable sera-t-elle disponible avant le cadrage, ou sa définition fait-elle partie du cadrage lui-même ? D'autre part, le projet parallèle de résorption du retard d'archivage (démarrant en septembre 2026) doit fournir à DataQuest un corpus préparé via COMED : quel est le calendrier prévu, et quel degré de confiance la CHD a-t-elle quant à la disponibilité de ce corpus avant le démarrage de DataQuest le 11 janvier 2027 ?

Réponse / Answer

FR : La CHD est responsable de la définition du plan de classement de ses documents. Sa définition sera prise en charge par les archivistes et fournie au titulaire au début de la première incrémentation. Le projet de traitement des archives numériques est pris en charge par la CHD selon les dates annoncées dans le CCTP. La CHD sera en mesure de fournir des lots/corpus



traitées pour le projet Dataquest. Le marché porte toutefois sur la mise en place d'une solution opérationnelle, et non sur le traitement exhaustif de ce corpus pendant le projet. Le titulaire devra donc déterminer et justifier le volume de documents nécessaire à la conception, à l'entraînement et à la validation de la solution, afin d'atteindre les niveaux de qualité attendus et de limiter les erreurs. Ce volume pourra porter sur tout ou partie du corpus, selon les besoins démontrés. À l'issue du projet, la CHD devra pouvoir traiter le solde des documents de manière autonome au moyen de la solution mise en place.

EN : The CHD is responsible for defining the classification scheme for its documents. This will be drawn up by the archivists and provided to the project lead at the start of the first iteration. The digital archives processing project is being managed by the CHD in accordance with the dates set out in the technical specifications. The CHD will be able to provide processed batches/corpora for the Dataquest project. However, the contract relates to the implementation of an operational solution, and not to the exhaustive processing of this corpus during the project. The contractor must therefore determine and justify the volume of documents required for the design, training and validation of the solution, in order to achieve the expected quality levels and minimise errors. This volume may cover all or part of the corpus, depending on the demonstrated requirements. Upon completion of the project, the CHD must be able to process the remaining documents independently using the solution put in place.

QUESTION 47

Le coût de l'hébergement est-il compris dans les 140 000 € hors taxes, ou pris en charge par la Chambre ou par l'État ?

Réponse / Answer

FR : Les coûts d'infrastructure et hébergement ne sont pas inclus dans le montant maximal de 140 000 EUR hors TVA du présent marché. Ils seront pris en charge séparément par la CHD. Le soumissionnaire doit néanmoins présenter une estimation détaillée des coûts de l'infrastructure cible nécessaire au fonctionnement de sa solution. Cette estimation ne sera pas intégrée au prix de l'offre, mais sera prise en compte dans son évaluation afin d'apprécier le coût global de la solution.

EN : Infrastructure and hostings costs are not included in the maximum amount of 140,000 EUR (excluding VAT) for this contract. They will be covered separately by the CHD. The bidder must nevertheless submit a detailed estimate of the costs of the target infrastructure required for the operation of its solution. This estimate will not be included in the bid price but will be taken into account in the evaluation to assess the overall cost of the solution.

QUESTION 48

Qui fournit les documents de test et les bonnes réponses de référence servant à mesurer la qualité de l'IA ? La Chambre les a-t-elle déjà, et combien de documents représentent-ils ?

Réponse / Answer



FR : Le titulaire devra constituer les ensembles de données en collaboration avec les archivistes de la CHD, comme évoqué en 4.3.2.

EN : The post holder will be required to compile the datasets in collaboration with the CHD archivists, as outlined in 4.3.2.

QUESTION 49

Parmi les neuf catégories de données sensibles, lesquelles sont considérées comme les plus graves, celles soumises au seuil de 99,5 % ?

Réponse / Answer

FR : Les neuf catégories du CCTP décrivent des types de sensibilité. Elles ne constituent pas un classement par gravité. La CHD devra arrêter et communiquer cette liste, avec les définitions applicables, avant l'évaluation des catégories concernées. C'est un travail qui doit être réalisé durant la phase de cadrage

EN : The nine categories in the technical specifications describe types of sensitive information. They are not ranked by severity. The CHD must finalise and communicate the list of critical categories, together with the applicable definitions, before those categories are evaluated. This work must be carried out during the scoping phase.

QUESTION 50

Les niveaux de qualité s'appliquent-ils seulement aux documents que l'outil a pu traiter, ou aussi à ceux écartés au départ parce qu'illisibles, corrompus ou dans un format non géré ?

Réponse / Answer

FR : Les niveaux de qualité s'appliquent aux documents que l'outil a pu traiter.

EN : The quality levels apply to the documents that the tool has been able to process.

QUESTION 51

Où en sera la réécriture du plan de classement au 11 janvier 2027, et sur quels dossiers portera la première étape du projet ?

Réponse / Answer

FR : Le fonds documentaire de la CHD est décrit au chapitre 3.3 du CCTP. Le travail sur le plan de classement sera défini progressivement, au cas par cas, pour chaque lot à traiter. La CHD fournira le plan de classement applicable, tandis que le titulaire devra impérativement proposer des lots de documents pertinents permettant d'itérer et d'améliorer la solution.

EN : The CHD's document collection is described in Chapter 3.3 of the Technical Specifications. Work on the classification scheme will be defined progressively, on a case-by-case basis, for each batch to be processed. The CHD will provide the applicable classification scheme, whilst the contractor must propose batches of relevant documents to enable the solution to be refined and improved.



QUESTION 52

Un hébergement chez un opérateur souverain comme LuxProvide (MeluXina) est-il recevable au titre du troisième scénario ? Et le service Mistral est-il accessible via le CTIE, à quelles conditions de volume et de facturation ?

Réponse / Answer

FR : L'ensemble des conditions pour l'hébergement est décrit dans le CCTP. Le soumissionnaire devra détailler l'ensemble des licences, souscriptions et droits d'usage et coûts de mise en œuvre et maintenance nécessaires au fonctionnement de la solution. Cette description devra couvrir notamment la plateforme principale, les connecteurs, les modèles d'IA, les moteurs OCR ainsi que tout composant tiers requis (chapitre 5.6.1).

EN : All the terms and conditions relating to hosting are set out in the Technical Specifications. The tenderer must provide details of all licences, subscriptions and usage rights, as well as the implementation and maintenance costs required for the solution to operate. This description must cover, in particular, the main platform, connectors, AI models, OCR engines and any required third-party components (section 5.6.1).

QUESTION 53

Une première version utilisable est attendue mi-février 2027 alors que le projet démarre le 11 janvier. Que doit-elle contenir au minimum, et la Chambre confirme-t-elle que les documents et les interlocuteurs seront disponibles dès la première semaine ?

Réponse / Answer

FR : Chaque incrément devra produire un résultat démontrable, évaluable et potentiellement exploitable. La mise en œuvre de la solution devra suivre une logique de progression maîtrisée, fondée sur l'intégration successive de périmètres documentaires de complexité croissante. Les premiers lots devront privilégier les fonds les plus simples à traiter afin d'obtenir rapidement des résultats mesurables et d'ajuster les paramétrages de la solution. La CHD mettra à disposition des lots documentaires et des interlocuteurs dès la première semaine du projet. Le premier périmètre attendu à mi-février 2027 est un lot limité de documents numériques textuels homogènes : un format, une langue, un système d'écriture et un périmètre métier cohérent. Il devra permettre de produire un inventaire démontrable et évaluable. La détection des données sensibles pourra se limiter à des cas élémentaires dans un premier temps.

EN : Each phase must produce a demonstrable, measurable and potentially usable result. The implementation of the solution must follow a controlled, step-by-step approach, based on the successive integration of document sets of increasing complexity. The first batches must prioritise the collections that are simplest to process in order to achieve measurable results quickly and to fine-tune the solution's settings. The CHD will provide document batches and points of contact from the first week of the project. The initial scope expected by mid-February 2027 is a limited batch of homogeneous digital text documents: one format, one language, one writing system and a coherent business domain. It must enable the production of an inventory



that can be demonstrated and evaluated. Sensitive data detection may initially be limited to basic cases.

QUESTION 54

Le cahier des charges liste huit types de problèmes de qualité sans donner de proportion. Quelle part du fonds est concernée, même en ordre de grandeur ? Un diagnostic a-t-il déjà été fait ? Et le nettoyage préalable de ces documents relève-t-il de la Chambre ou du prestataire ?

Réponse / Answer

FR : Les fonds sont en principe déjà classés et triés. Le titulaire devra néanmoins être en capacité de mettre en œuvre, en accord avec la CHD et selon les besoins constatés, des traitements préparatoires tels que le nettoyage, la normalisation, la restructuration ou l'exclusion de certains documents. Aucun ordre de grandeur n'est communiqué à ce stade.

EN : In principle, the records have already been classified and sorted. The holder must, however, be able to carry out, in consultation with the CHD and in accordance with identified needs, preparatory processes such as cleaning, standardisation, restructuring or the exclusion of certain documents. No estimate of the scale of the work is provided at this stage.

QUESTION 55

Quel est l'objectif réel de fin de projet : avoir traité les 4,3 millions de documents du fonds, ou avoir démontré que la solution fonctionne sur les fonds sélectionnés pour les phases pilotes ? Un volume cible engageant est-il attendu dans l'offre ?

Réponse / Answer

FR : Les objectifs de réussite sont décrits dans le chapitre 2.3.2 : La réussite du projet sera appréciée au regard de la capacité de la solution à produire des résultats exploitables dans les délais impartis, afin de démontrer rapidement sa faisabilité et sa valeur ajoutée pour les équipes de la CHD. Le marché porte toutefois sur la mise en place d'une solution opérationnelle, et non sur le traitement exhaustif de ce corpus pendant le projet. Le titulaire devra donc déterminer et justifier le volume de documents nécessaire à la conception, à l'entraînement et à la validation de la solution, afin d'atteindre les niveaux de qualité attendus et de limiter les erreurs. Ce volume pourra porter sur tout ou partie du corpus, selon les besoins démontrés. À l'issue du projet, la CHD devra pouvoir traiter le solde des documents de manière autonome au moyen de la solution mise en place.

EN : The success criteria are set out in section 2.3.2: The success of the project will be assessed on the basis of the solution's ability to produce usable results within the specified timeframe, in order to demonstrate its feasibility and added value to the CHD teams at an early stage. However, the contract relates to the implementation of an operational solution, and not to the exhaustive processing of this corpus during the project. The contractor must therefore determine and justify the volume of documents required for the design, training and validation of the solution, in order to achieve the expected quality levels and minimise errors. This volume may cover all or part of the corpus, depending on the demonstrated requirements. Upon completion of the project, the



CHD must be able to process the remaining documents independently using the solution put in place.

QUESTION 56

Quelle est la taille réelle des documents du fonds : nombre moyen et maximum de pages, poids moyen et maximum des fichiers, et part des documents très volumineux ? La liste des formats de référence inclut MP3, MP4, WAV et SIARD : des fichiers audio, vidéo et des exports de bases de données sont-ils effectivement présents dans le périmètre de la V1, et avec quelles durées ou volumes ?

Réponse / Answer

FR : Le pouvoir adjudicateur ne dispose pas, à ce stade, d'informations suffisamment consolidées pour fournir une répartition fiable du corpus par type de document, format, taille, volume, taux de numérisation ou d'OCR, ou niveau de standardisation. Les soumissionnaires construiront leur offre sur la base des informations quantitatives et qualitatives figurant dans les documents du marché. Dans leur offre, ils devront préciser et justifier les types, formats et volumes de documents ainsi que les autres caractéristiques du corpus nécessaires pour assurer l'efficacité et la qualité de leur solution.

EN : At this stage, the contracting authority does not have sufficiently consolidated information to provide a reliable breakdown of the corpus by document type, format, size, volume, digitisation or OCR rate, or level of standardisation. Tenderers shall base their tenders on the quantitative and qualitative information contained in the tender documents. In their tenders, they must specify and justify the types, formats and volumes of documents, as well as any other characteristics of the corpus necessary to ensure the effectiveness and quality of their solution.

QUESTION 57

Quelle est la répartition des documents par langue, notamment la part du luxembourgeois ? Les modèles d'IA sont moins performants en luxembourgeois qu'en français, en allemand ou en anglais. Un traitement différencié est-il envisagé pour cette langue, ou un seuil ajusté ?

Réponse / Answer

FR : Le pouvoir adjudicateur ne dispose pas, à ce stade, d'informations suffisamment consolidées permettant de fournir une répartition fiable du corpus selon les langues de documents. L'exigence est similaire sur le traitement du luxembourgeois.

EN : At this stage, the contracting authority does not have sufficiently consolidated information to provide a reliable breakdown of the corpus by document language. The requirement is similar with regard to the treatment of the Luxembourgish language.

QUESTION 58

Afin de mieux dimensionner notre approche technique et le chiffrage de notre offre, serait-il possible d'accéder, avant la remise des offres, à un échantillon représentatif (anonymisé ou non



sensible) du fonds documentaire de la CHD, compte tenu de l'hétérogénéité des formats et des langues mentionnée dans le cahier des charges ?

Réponse / Answer

FR : Aucun échantillon représentatif du corpus ne sera mis à la disposition des soumissionnaires avant la date limite de dépôt des offres.

EN : A representative sample of the corpus will not be made available to tenderers before the submission deadline.

QUESTION 59

Dans le scénario d'un hébergement assuré par le CTIE, les capacités de calcul (GPU) et de stockage nécessaires aux traitements IA sont-elles déjà mutualisées et mises à disposition sans coût additionnel pour le titulaire, ou doivent-elles être intégralement dimensionnées et chiffrées par le titulaire dans son offre, même si hébergées chez le CTIE ?

Réponse / Answer

FR : Les coûts d'infrastructure et hébergement ne sont pas inclus dans le montant maximal de 140 000 EUR hors TVA du présent marché. Ils seront pris en charge séparément par la CHD.

Le soumissionnaire doit néanmoins présenter une estimation détaillée des coûts de l'infrastructure cible nécessaire au fonctionnement de sa solution. Cette estimation ne sera pas intégrée au prix de l'offre, mais sera prise en compte dans son évaluation afin d'apprécier le coût global de la solution.

EN : Infrastructure and hostings costs are not included in the maximum amount of 140,000 EUR (excluding VAT) for this contract. They will be covered separately by the CHD. The bidder must nevertheless submit a detailed estimate of the costs of the target infrastructure required for the operation of its solution. This estimate will not be included in the bid price but will be taken into account in the evaluation to assess the overall cost of the solution.

QUESTION 60

Participation et références - CCTP 1.5, 1.13, 1.14, 8.6 et 8.7.1

Un opérateur non invité nominativement peut-il candidater ? Pour les trois références requises, les produits internes sont-ils admissibles, notamment Medelu, notre logiciel pour le secteur pharmaceutique, pour ses travaux de qualité des données, de rapprochement et de traçabilité ? Quels justificatifs de réalisation et de comparabilité sont attendus ? Les expériences des intervenants chez leurs employeurs peuvent-elles être retenues, avec attribution exacte des rôles ? En groupement, ce critère est-il collectif ou individuel ? Les références d'un sous-traitant engagé sont-elles mobilisables, et doit-il lui-même satisfaire au critère ?

Réponse / Answer

FR : Il s'agit d'une procédure restreinte sans publication d'avis pour un marché public de faible envergure (Livre I de la Loi du 8 avril 2018 sur les marchés publics). Un opérateur économique



peut envisager de soumettre une offre de candidature en ayant recours à la sous-traitance. La part de marché et les activités que l'opérateur économique entend sous-traiter devra être indiquée dans son offre (chapitre 1.13). Conformément au chapitre 8.6, le candidat devra justifier d'au moins trois références de projets similaires. Les projets présentés décriront le même type de technologie. Ils porteront soit sur l'analyse de corpus documentaires hétérogènes soit sur des organismes publics. L'ensemble des opérateurs économiques de l'offre collective devront répondre aux critères de sélection sous peine d'irrecevabilité de celle-ci (cf.1.14).

EN : This is a restricted procedure without publication of a notice for a small-scale public contract (Book I of the Public Procurement Act of 8 April 2018). An economic operator may consider submitting a tender by making use of subcontractors. The proportion of the contract and the activities that the economic operator intends to subcontract must be specified in its tender (Chapter 1.13). In accordance with Chapter 8.6, the tenderer must provide evidence of at least three references for similar projects. The projects presented must describe the same type of technology. They must relate either to the analysis of heterogeneous document corpora or to public bodies. All economic operators participating in the joint tender must meet the selection criteria, failing which the tender will be deemed inadmissible (1.14).

QUESTION 61

Profils et justificatifs - CCTP 8.7.2, 8.7.3 et 9.3 :

Les exigences de diplôme, d'expérience et de langue concernent-elles les responsables fonctionnel et technique évalués ou chaque intervenant ? Quels justificatifs sont admis pour l'expérience entrepreneuriale et la langue maternelle ? À quelle date les quatre années d'expérience doivent-elles être acquises, et quels éléments sont requis dès l'offre pour les remplaçants ?

Réponse / Answer

FR : Les exigences relatives au diplôme, à l'expérience et à la langue s'appliquent à l'ensemble des intervenants. L'expérience entrepreneuriale peut être attestée par une simple description accompagnée des coordonnées des entreprises ou institutions citées. Aucun justificatif n'est requis pour la langue maternelle : une simple mention suffit. Les quatre années d'expérience devront être acquises au plus tard au 31 décembre 2026. Enfin, les remplaçants devront faire l'objet, dès l'offre, des mêmes informations et éléments que les titulaires.

EN : The requirements regarding qualifications, experience and language apply to all applicants. Entrepreneurial experience may be demonstrated by a brief description accompanied by the contact details of the companies or institutions mentioned. No supporting documentation is required for the mother tongue: a simple statement is sufficient. The four years' experience must have been gained by 31 December 2026 at the latest. Finally, stand-ins must be provided with the same information and details as the permanent staff members from the time the offer is made.

QUESTION 62

Corpus et fonctions à livrer - CCTP 2.2, 3.3, 3.4.4, 4.1.4 et 5.1.7 :



Quel volume doit être effectivement traité pendant le marché, distinctement de la capacité attendue ? Un échantillon représentatif, la matrice d'inventaire et la répartition par langue, format, taille et qualité de numérisation seront-ils accessibles aux candidats ? Pour SIARD, les contenus audio et vidéo, attend-on les métadonnées seules ou l'extraction et la transcription intégrales ? La « génération automatisée de réponses » implique-t-elle une interface de questions-réponses supplémentaire ?

Réponse / Answer

FR : Aucun document supplémentaire n'est fourni par la CHD durant cette étape. Les attendus concernant les contenus audio et vidéo sont les mêmes que ceux pour les documents, ils sont décrits au chapitre 4.2.3. Le terme « génération automatisée de réponses », mentionné à la section 2.2, fait référence à la restitution en langage naturel au sein de la matrice d'inventaire décrite à la section 3.2.4, étape 7.

EN : No additional documents are provided by the CHD at this stage. The requirements regarding audio and video content are the same as those for documents; they are described in section 4.2.3. The term "génération automatisée de réponses", mentioned in section 2.2, refers to the output in natural language within the inventory matrix described in section 3.2.4, step 7.

QUESTION 63

Infrastructure et coûts de traitement - CCTP 1.13, 1.18, 5.1.3 à 5.1.6, 5.6.1 et 9.2 :

Au plafond de 140 000 EUR HT, nous retenons 126 000 EUR de prestations initiales et 14 000 EUR de contingence. Pour les quatre environnements, quelles dépenses sont financées par la CHD / le CTIE et lesquelles sont à inclure au forfait : calcul, GPU, stockage, sauvegardes, supervision, licences, OCR et inférence, y compris tests, reprises et hypercare ? Quels modèles, quotas et tarifs de référence seront fournis ? Les achats de ces services entrent-ils dans le plafond de 45 % de sous-traitance, et selon quelle assiette ?

Réponse / Answer

FR : Pour l'aspect budgétaire, merci de se référer à la réponse faite à la question 8. Aucun modèle, quotas et tarifs de référence n'est fourni. L'ensemble des éléments sont présents dans le CCTP.

EN : With regard to the budgetary aspect, please refer to the answer to question 8. No templates, quotas or reference rates are provided. All relevant details are set out in the Specifications.

QUESTION 64

Périmètre autorisé et assistance par IA - CCTP 3.2.2.5, 5.1.3, 5.1.4, 5.2.1 et 9.7 :

Comment concilier les environnements souverains ou autorisés envisagés avec l'interdiction des transferts vers des services de cloud privés en section 9.7 ? Quels modèles, services d'inférence et outils d'assistance au développement sont autorisables, avec quelles exigences de localisation, conservation, accès et traçabilité ? Merci de distinguer les usages sur données publiques ou synthétiques sans contenu confidentiel de ceux portant sur les données, le code ou les éléments confidentiels de la CHD.

Réponse / Answer



FR : Les données de la CHD et leurs traitements doivent rester dans des infrastructures validées par elle. La section 9.7 interdit leur transfert vers un cloud privé. La localisation européenne ou la qualification commerciale de « souverain » ne suffit pas à lever cette interdiction. Aucun outil externe d'assistance au développement ne doit recevoir des données, du code non public ou des éléments confidentiels du projet. Les modèles, flux, lieux de traitement, durées de conservation, accès et journaux devront être décrits et validés au cadrage. Le soumissionnaire devra identifier les services proposés et préciser, pour chacun : la localisation des traitements et du stockage, les sous-traitants, les durées de conservation, l'absence de réutilisation à des fins d'entraînement, les modalités d'effacement, le contrôle des accès et la traçabilité des opérations.

EN : CHD data and its processing must remain within infrastructure approved by the CHD. Section 9.7 prohibits transferring that data to a private cloud. Hosting within Europe or describing a service as "sovereign" is not sufficient to override this prohibition. No external development assistant may receive project data, non-public code or confidential information. The models, data flows, processing locations, retention periods, access controls and logs must be described and approved during the scoping phase. The tenderer must identify the services offered and specify, for each one: the location of processing and storage, any data processors, retention periods, the absence of reuse for training purposes, the procedures for erasure, access control and the traceability of operations.

QUESTION 65

Mesure de la qualité et validation métier - CCTP 6.4.6 :

Les minima chiffrés étant pris en compte, quelles unités d'évaluation, catégories critiques et tailles minimales de strates seront retenues ? Comment les abstentions, documents inexploitable et renvois en revue humaine entrent-ils dans les dénominateurs, avant correction humaine ? Comment sont évaluées les catégories rares et les incertitudes statistiques ? Qui fournit et annote les jeux de référence, arbitre les désaccords et assure la disponibilité des archivistes pour la validation ?

Réponse / Answer

FR : Les unités d'évaluation seront définies en fonction de chaque indicateur (champ complété, valeur structurée, notice descriptive, preuve associée, etc.). Pour chaque opération évaluée, le taux de conformité est égal au nombre d'unités conformes divisé par le nombre total d'unités contrôlées, multiplié par 100. Les abstentions, documents inexploitable, et renvois en revue humaine sont traités en étape initiale de la chaîne de traitement (chapitre 4.1.1). Les minima fixés évaluent les lots soumis au traitement automatisés d'IA. Les cas critiques ou litigieux feront l'objet d'une double revue métier. L'équipe projet de la CHD est définie au chapitre 2.1.

EN : Assessment units will be defined according to each indicator (completed field, structured value, descriptive note, associated evidence, etc.). For each operation assessed, the compliance rate is equal to the number of compliant units divided by the total number of units checked, multiplied by 100. Abstentions, unusable documents and cases referred for human review are handled at the initial stage of the processing chain (section 4.1.1). The minimum thresholds set apply to batches subjected to automated AI processing. Critical or disputed cases will be subject to a double review by subject matter experts. The CHD project team is defined in section 2.1.



QUESTION 66

Essai de charge et capacité garantie - CCTP 5.5.1 et 6.4.6.9 :

Pour l'essai de 100 000 documents en huit heures, quel profil de corpus et quelle charge simultanée font référence ? Quelles ressources et quels quotas seront garantis ? Comment les indisponibilités ou limitations d'un service imposé par la CHD / le CTIE seront-elles prises en compte dans l'évaluation ?

Réponse / Answer

FR : Aucune information complémentaire concernant le profil de corpus et la charge simultanée ne sera fournie par la CHD durant la phase d'offre. L'ensemble des informations sont présentes dans le chapitre 3.3 Données et fonds documentaires. La CHD prendra en compte la cohérence de l'offre proposée par rapport aux besoins fonctionnels et techniques du CCTP, ainsi que les justifications données. Le soumissionnaire doit proposer une offre cohérente sur ce point.

EN : No further information regarding the corpus profile and concurrent load will be provided by the CHD during the tender phase. All relevant information is set out in Chapter 3.3: Data and Document Collections. The CHD will assess the consistency of the tender proposed with the functional and technical requirements of the Specifications, as well as the justifications provided. The tenderer must submit a tender that is consistent in this respect.

QUESTION 67

Prérequis, calendrier et admissions - CCTP 1.17, 6.3 à 6.6 et 9.2 :

Pour le démarrage du 11 janvier 2027, le premier périmètre à mi-février et les quatre incréments, quelles dates sont prévues pour les accès, environnements, modèles et corpus ? Quels délais de validation métier, juridique et SSI sont retenus ? Quel mécanisme encadre les effets d'un retard d'un prérequis ou d'une décision côté CHD / CTIE sur les délais, admissions, paiements et coûts ?

Réponse / Answer

FR : Les dates de mise à disposition des prérequis et les délais de validation métier, juridique et SSI seront précisés d'un commun accord lors du cadrage, sur la base du calendrier contractuel. La CHD s'engage à respecter les échéances convenues pour les actions relevant de sa responsabilité. Tout retard affectant un prérequis ou une décision doit être signalé sans délai et documenté. Le titulaire précise son incidence effective sur les échéances concernées et propose les mesures permettant d'en limiter les conséquences. Les impacts et les mesures à retenir sont examinés dans les instances de gouvernance et formalisés par écrit. Le titulaire ne pourra être sanctionné, au titre du respect des délais, pour la part du décalage directement causée par un retard dont il est établi qu'il ne lui est pas imputable. Il reste tenu de ses obligations d'anticipation, d'alerte et de limitation des conséquences du retard. Les admissions et paiements restent soumis aux conditions prévues aux sections 6.4.8 et 9.2. Les éventuelles incidences financières feront l'objet d'un examen distinct dans le respect du cadre contractuel et légal applicable.

EN : The dates by which prerequisites are to be made available and the deadlines for business, legal and information security validation will be specified by mutual agreement during the scoping phase, based on the contractual timetable. The CHD undertakes to meet the agreed



deadlines for actions falling within its remit. Any delay affecting a prerequisite or a decision must be reported without delay and documented. The contractor shall specify the actual impact on the deadlines concerned and propose measures to minimise the consequences. The impacts and the measures to be adopted shall be examined by the governance bodies and formalised in writing. The contractor shall not be penalised, in terms of compliance with deadlines, for the portion of the delay directly caused by a delay which is established as not being attributable to it. The contractor remains bound by its obligations to anticipate, give notice of and minimise the consequences of the delay. Admissions and payments remain subject to the conditions set out in sections 6.4.8 and 9.2. Any financial implications will be examined separately in accordance with the applicable contractual and legal framework.

QUESTION 68

Hypercare et partage des interventions - CCTP 7.1 à 7.3 :

Quelles sont les plages ouvrées de la CHD et les responsabilités respectives du titulaire, du CTIE et des fournisseurs pour le diagnostic et le rétablissement ? Une correction à périmètre inchangé relance-t-elle trois mois d'hypercare ? Comment sont traités les incidents résiduels à l'échéance maximale de onze mois ?

Réponse / Answer

FR : Le périmètre de l'Hypercare est défini au chapitre 7.3.1. Les pages ouvrées de la CHD seront définies au cadrage, comme mentionné dans le chapitre 7.2.1. La cloture d'Hypercare suppose l'absence d'incident critique ouvert, la résolution ou l'acceptation formelle des autres anomalies, le respect des seuils de qualité sur une période de quatre semaines, la réussite des tests de non-régression et la capacité de la CHD à administrer le périmètre.

EN : The scope of Hypercare is defined in section 7.3.1. The working pages of the CHD will be defined during the scoping phase, as mentioned in section 7.2.1. The closure of Hypercare requires that there are no open critical incidents, that all other issues have been resolved or formally accepted, that quality thresholds have been met over a four-week period, that non-regression tests have been successfully completed, and that the CHD is capable of managing the scope.

QUESTION 69

Évolutions et contingence - CCTP 1.18, 7.3.2 et 9.6 :

Comment articuler la contingence pour changements approuvés avec l'interdiction de modifier le marché en section 9.6 ? Quelle procédure distingue une correction incluse d'un nouveau besoin ou d'une évolution de règle métier validée, et autorise ses conséquences sur le prix et le calendrier ?

Réponse / Answer

FR : Le budget de contingence, fixé à 10 % du montant du marché, constitue un montant maximal destiné à financer d'éventuelles prestations supplémentaires. Son utilisation est subordonnée à une demande ou à une autorisation écrite préalable du pouvoir adjudicateur, sur la base d'une estimation détaillée du titulaire, préalablement validée par celui-ci. Sont considérées comme



supplémentaires les prestations qui ne figurent pas dans l'expression des besoins ni dans le périmètre contractuel initial. Les adaptations et évolutions des règles métier prévues dans le périmètre du marché sont comprises dans le prix initial et ne peuvent donner lieu à la mobilisation du budget de contingence.

EN : The contingency budget, set at 10 per cent of the contract value, represents a maximum amount intended to finance any additional services. Its use is subject to a request or prior written authorisation from the contracting authority, based on a detailed estimate from the contractor, which has been previously approved by the contracting authority. Services that do not appear in the statement of requirements or within the initial scope of the contract are considered to be additional. Any adjustments or changes to the business rules provided for within the scope of the contract are included in the initial price and may not give rise to the use of the contingency budget.

QUESTION 70

Responsabilité et assurance - CCTP 9.5, 9.10 ; annexe 10.2, art. 8 :

Un plafond global de responsabilité est-il prévu ? Quelles garanties et quels montants d'assurance sont exigés ? Pour la pénalité de confidentialité de 2 500 EUR par violation, quelle unité de décompte s'applique et comment s'articule-t-elle avec d'autres indemnisations du même fait ?

Réponse / Answer

FR : Le CCTP ne prévoit pas de plafond global contractuel de responsabilité. La responsabilité du titulaire est régie par la section 9.10 et par le droit luxembourgeois. Le titulaire doit respecter les obligations relatives à l'assurance de responsabilité professionnelle visées à la section 9.5. Aucun montant minimal de couverture ni garantie d'assurance particulière supplémentaire n'est fixé par le CCTP. Les limites de couverture de son assurance ne constituent pas un plafond de sa responsabilité envers la CHD. La pénalité de 2 500 EUR prévue à l'article 8 de l'accord de confidentialité s'applique à chaque violation distincte et établie de cet accord. Le décompte repose sur les manquements constatés ; le nombre de documents, de données ou de personnes concernés ne constitue pas, à lui seul, une unité automatique de décompte. Cette pénalité est soumise au régime des clauses pénales en droit luxembourgeois, notamment à l'article 1152 du Code civil. Elle indemnise forfaitairement le manquement qu'elle sanctionne et ne se cumule pas avec des dommages-intérêts réparant le même préjudice couvert par la clause. Elle ne constitue pas un plafond global de responsabilité du titulaire et n'emporte aucune renonciation de la CHD aux autres droits et recours prévus par le CCTP et le droit applicable, notamment les demandes relatives à des préjudices distincts non couverts par cette pénalité et la résiliation du marché lorsque ses conditions sont réunies. Le pouvoir du juge de réduire ou d'augmenter une pénalité manifestement excessive ou dérisoire demeure réservé.

EN : The Tender Documents do not provide for an overall contractual limit of liability. The contractor's liability is governed by Section 9.10 and by Luxembourg law. The contractor must comply with the obligations relating to professional indemnity insurance set out in Section 9.5. No minimum level of cover or specific additional insurance cover is stipulated in the Tender Documents. The limits of cover under the contractor's insurance do not constitute a cap on the



contractor's liability towards the CHD. The penalty of EUR 2,500 provided for in Article 8 of the confidentiality agreement applies to each separate and established breach of that agreement. The calculation is based on the breaches found; the number of documents, data items or persons concerned does not, in itself, constitute an automatic unit of calculation. This penalty is subject to the provisions governing penalty clauses under Luxembourg law, in particular Article 1152 of the Civil Code. It provides lump-sum compensation for the breach it penalises and cannot be combined with damages awarded for the same loss covered by the clause. It does not constitute an overall cap on the contractor's liability and does not imply any waiver by CHD of the other rights and remedies provided for in the Special Conditions of Contract and applicable law, in particular claims relating to separate losses not covered by this penalty and the termination of the contract where the conditions for such termination are met. The court retains the power to reduce or increase a penalty that is manifestly excessive or derisory.

QUESTION 71

Composants antérieurs et droits - CCTP 9.8

Quelle procédure permet de déclarer et de faire valider les composants antérieurs, tiers et spécifiques au marché, ainsi que leurs droits d'usage ? Comment sont traitées les adaptations d'éléments antérieurs et la mutualisation par le CTIE ? Quels justificatifs sont attendus pour un démonstrateur constitué avant attribution ?

Réponse / Answer

FR : La section 9.8 du CCTP définit le régime applicable. Les composants préexistants du titulaire, les composants tiers et les développements spécifiques peuvent être présentés dans un inventaire précisant les titulaires des droits, les licences applicables et les éventuelles restrictions. Les composants tiers doivent être identifiés et communiqués à la CHD et faire l'objet de son accord préalable et écrit avant leur utilisation. Les licences applicables doivent être compatibles avec les exigences d'exploitation, de maintenance, de réversibilité et de mutualisation prévues au CCTP. L'accord de la CHD ne décharge pas le titulaire de son obligation de garantir qu'il dispose de tous les droits et autorisations nécessaires aux cessions et licences prévues à la section 9.8. Les éléments préexistants restent la propriété de leurs titulaires, avec octroi à la CHD des droits nécessaires prévus à la section 9.8. Les adaptations protégeables réalisées spécifiquement dans le cadre du marché relèvent de la cession prévue à cette section, sans transfert de propriété des éléments préexistants eux-mêmes. La mutualisation par l'intermédiaire du CTIE concerne le socle technique et les composants génériques réutilisables, à l'exclusion des données, documents, corpus, métadonnées et résultats métier propres à la CHD. Le CCTP ne prescrit pas de forme particulière de justificatif pour un démonstrateur constitué avant l'attribution. Si ses composants sont intégrés à la solution, le titulaire doit pouvoir justifier leur antériorité et les droits permettant leur utilisation, notamment par une documentation datée, un historique de versions et les licences correspondantes. Les éléments préexistants doivent être distingués des développements spécifiques réalisés dans le cadre du marché. La présentation du démonstrateur ne vaut ni validation des droits sur ses composants ni dérogation aux obligations de la section 9.8.



EN : Section 9.8 of the Technical Specifications defines the applicable regime. The contractor's pre-existing components, third-party components and specific developments may be set out in an inventory specifying the rights holders, the applicable licences and any restrictions. Third-party components must be identified and notified to the CHD, and must be subject to the CHD's prior written consent before they are used. The applicable licences must be compatible with the requirements for operation, maintenance, reversibility and pooling set out in the Tender Documents. The CHD's approval does not relieve the contractor of its obligation to ensure that it holds all the necessary rights and authorisations for the transfers and licences provided for in Section 9.8. Pre-existing elements remain the property of their owners, with the CHD being granted the necessary rights as set out in section 9.8. Protectable adaptations made specifically within the scope of the contract are subject to the assignment provided for in that section, without any transfer of ownership of the pre-existing elements themselves. The Technical Specifications do not stipulate any particular form of supporting documentation for a demonstrator built prior to the award of the contract. If its components are incorporated into the solution, the contractor must be able to demonstrate their prior existence and the rights permitting their use, in particular by means of dated documentation, a version history and the relevant licences. Pre-existing elements must be distinguished from specific developments carried out under the contract. The presentation of the demonstrator does not constitute validation of the rights to its components nor an exemption from the obligations set out in Section 9.8.

QUESTION 72

Accords, conservation et incidents de données - CCTP 4.3.2, 5.3.1, 6.4.6 ; annexes 10.2 à 10.4 :

- Comment l'autorisation préalable des sous-traitants de données s'articule-t-elle avec le délai d'objection de trois mois ? Une acceptation expresse permet-elle leur mobilisation au démarrage ?
- La notification d'une violation sous deux heures s'applique-t-elle hors plages ouvrées, et selon quel dispositif ?
- Quelle matrice de conservation concilie la suppression des copies avec les preuves, jeux de référence, journaux et sauvegardes requis ?
- Les mentions manquantes aux articles 4 à 6 du NDA société seront-elles corrigées ? Quel modèle individuel convient aux dirigeants et indépendants, et quand les accords doivent-ils être signés ?

Réponse / Answer

FR : Le recours à un sous-traitant ultérieur est soumis à l'autorisation préalable et écrite de la CHD, conformément à l'article 28 du RGPD et aux accords annexés au CCTP. Le délai de trois mois correspond au délai maximal dont dispose la CHD pour présenter ses objections après réception de l'ensemble des informations prévues à l'annexe 10.4. Une autorisation expresse et écrite de la CHD peut intervenir avant son expiration et permettre le démarrage des activités à la date qu'elle précise, sans attendre l'écoulement des trois mois. Les sous-traitants identifiés dans l'offre peuvent ainsi être autorisés lors de la signature du contrat et intervenir dès le démarrage du projet, sous réserve de la mise en place des engagements contractuels requis. Leur seule mention dans l'offre ne vaut pas autorisation. Cette autorisation ne dispense pas le titulaire de vérifier les garanties présentées par le sous-traitant ultérieur et de lui imposer les obligations de



protection des données requises. Le titulaire demeure pleinement responsable envers la CHD de l'exécution de ces obligations par ses sous-traitants ultérieurs. La notification d'une violation sous deux heures s'applique également hors des plages ouvrées. La notification doit se faire par courriel, conformément au point 9 de l'article IV de l'accord relatif à la protection des données figurant en annexe 10.4. Les modalités d'escalade, de communication et de notification des parties concernées devront être documentées. La suppression porte sur les données et leurs copies, et non sur les éléments de journalisation nécessaires à la traçabilité. Les données sources, copies de travail et jeux de référence contenant des données du CHD, devront être supprimés à l'issue de leur durée d'utilisation autorisée. Les journaux et preuves d'exécution, seront conservés pendant la durée définie par le CHD, sans contenir le contenu des documents ni de données confidentielles non nécessaires. Les versions actuelles des accords de confidentialité sont complètes. Deux modèles distincts sont prévus : un modèle individuel applicable aux personnes physiques, notamment aux salariés, dirigeants et indépendants, et un modèle destiné aux entreprises. Les accords applicables devront être signés en même temps que le contrat.

EN : The use of a sub-contractor is subject to the prior written authorisation of the CHD, in accordance with Article 28 of the GDPR and the agreements annexed to the CCTP. The three-month period corresponds to the maximum time limit available to the CHD to raise any objections following receipt of all the information set out in Annex 10.4. Express written authorisation from the CHD may be granted before this period expires, allowing activities to commence on the date specified by the CHD, without waiting for the three months to elapse. The sub-processors identified in the tender may thus be authorised upon signing the contract and may commence work from the start of the project, subject to the required contractual commitments being put in place. Merely mentioning them in the tender does not constitute authorisation. This authorisation does not exempt the contractor from verifying the safeguards provided by the subsequent subcontractor and from imposing the required data protection obligations on them. The contractor remains fully liable to the CHD for the fulfilment of these obligations by its subsequent subcontractors. The requirement to report a breach within two hours also applies outside normal working hours. The report must be made by email, in accordance with paragraph 9 of Article IV of the data protection agreement set out in Annex 10.4. The procedures for escalation, communication and notification of the parties concerned must be documented. Deletion applies to the data and its copies, but not to the log entries required for traceability. Source data, working copies and reference sets containing CHD data must be deleted at the end of their authorised period of use. Logs and execution records shall be retained for the period defined by the CHD, without containing the content of the documents or any unnecessary confidential data. The current versions of the confidentiality agreements are complete. Two separate templates are provided: an individual template applicable to natural persons, in particular employees, directors and self-employed individuals, and a template intended for companies. The applicable agreements must be signed at the same time as the contract.

QUESTION 73

Dossier et démonstration - CCTP 8.7.4, 8.8 et 8.9



Au-delà des exclusions explicites à la limite de huit pages, des annexes techniques sont-elles autorisées et évaluées ? Les cinq points liés à une démonstration sont-ils inclus dans les quinze points de présentation ou additionnels ? Quel total maximal en résulte ?

Réponse / Answer

FR : Les cinq points associés à la démonstration facultative prévue à la section 8.7.4 sont des points bonus, qui s'ajoutent aux 100 points attribués au titre des cinq critères. Pour les annexes, se référer à la réponse de la question 24.

EN : The five points associated with the optional demonstration provided for in Section 8.7.4 are bonus points, awarded in addition to the 100 points available under the five criteria. For annex, please refer to the answer to question 24.