

# Appel à propositions : Ensembles de données pour les langues en Afrique subsaharienne

## Questions/réponses

Remarque: Ce document a été révisé le 28 octobre pour corriger une erreur à la question 11.

Q0	Qui seront les évaluateurs ? Sont-ils issus uniquement du GCT ou devons-nous désigner des personnes dans le formulaire de proposition ?
R0	Les membres du groupe consultatif technique, avec le soutien du comité directeur et d'autres partenaires, évalueront les propositions. Il n'est pas nécessaire de désigner des évaluateurs dans le formulaire de proposition.
Q1	Une proposition peut-elle porter sur plusieurs ensembles de données, par exemple, s'ils sont liés ?
R1	<p>Oui. La portée d'une proposition doit être scientifiquement cohérente et adaptée aux cas d'utilisation, en TALN et pour les applications en aval, qui motivent la création du ou des ensembles de données.</p> <p>Par exemple, une proposition comprenant des ensembles de données en plusieurs langues pour améliorer la précision d'une tâche spécifique de TALN dans divers contextes serait acceptable. De même, une collection d'ensembles de données pour plusieurs tâches de TALN dans une langue ou un contexte particulier, qui généreraient conjointement un cas d'utilisation ayant un impact concret, serait acceptable.</p> <p>Les candidats peuvent également soumettre plus d'une proposition dans le cas de plusieurs projets sans lien entre eux. Veuillez également consulter la question 11 sur les considérations additionnelles relatives à la soumission de propositions plus importantes. Le groupe consultatif technique peut choisir de financer toute proposition en tout ou en partie.</p>
Q2	Une idée de la taille des ensembles de données ?
R2	L'exposé de la proposition (voir la section 4 : Exigences de la proposition) demande aux candidats de présenter l'ensemble de données qu'ils proposent, y compris la taille de l'ensemble de données, dans le contexte des cas d'utilisation proposés pour ces données. Les propositions doivent garantir que la taille de l'ensemble de données proposé permet une analyse significative des phénomènes inclus dans l'ensemble de données et/ou une formation et une évaluation suffisantes pour les modèles d'apprentissage machine.
Q3	<p>1. Quel est le montant total du financement ?</p> <p>2. Combien de projets cherchez-vous à financer ?</p>

R3	Le nombre et le coût total des projets financés dépendront de la qualité des propositions. Le Fonds visera à financer un éventail de projets dans toutes les régions et à toutes les échelles.
Q4	Le candidat principal doit-il être un spécialiste des données ? Je suis éthicien en IA et gestionnaire de projets internationaux et je peux constituer une équipe capable de mener à bien le développement d'un ensemble de données. Puis-je tout de même proposer ma candidature ?
R4	Non, le candidat principal ne doit pas nécessairement être un spécialiste des données. L'équipe doit avoir l'expertise nécessaire pour mener à bien le projet. Pour plus de détails sur la manière dont les qualifications seront évaluées, voir les critères d'éligibilité dans l'appel à propositions et la rubrique « Qualité » parmi les critères d'évaluation.
Q5	Pouvez-vous donner plus de détails sur la question de l'approbation éthique ? Doit-elle être jointe à la proposition ?
R5	L'approbation éthique ne doit pas nécessairement être finalisée au moment de la soumission de la proposition. Les propositions doivent démontrer qu'elles tiennent compte des problèmes éthiques potentiels auxquels elles peuvent être confrontées, de la manière dont elles prévoient d'y faire face, et de tout processus institutionnel ou national qu'elles devront subir (voir la section 4 : Exigences de la proposition). Dans le cadre du processus d'évaluation, le groupe consultatif technique peut signaler les problèmes éthiques à résoudre si la subvention est accordée.
Q6	Pouvons-nous recueillir des données auprès d'une seule source ? Dans une communauté, il peut y avoir plusieurs langues. Dans ce cas, que peut-on faire ?
R6	Les sources et les méthodes de collecte de données doivent être techniquement solides et correspondre aux cas d'utilisation de l'apprentissage machine que l'ensemble de données proposé générera. Les sources et méthodes de collecte doivent également garantir que les données sont représentatives et, dans la mesure du possible, impartiales. Voir la question 1 pour plus de détails sur les propositions portant sur plusieurs langues ou sur d'autres ensembles de données connexes.
Q7	Des propositions multiples (pas nécessairement en TALN) peuvent-elles être soumises par la même institution de recherche ?
R7	Oui. Toutefois, veuillez noter que cet appel à propositions concerne uniquement les ensembles de données de formation et d'évaluation pour les langues africaines.  Consultez le <a href="#">site web</a> du Lacuna Fund et abonnez-vous à notre newsletter pour rester informé des futures possibilités de financement dans d'autres domaines.
Q8	Des institutions peuvent-elles soumettre une demande conjointe ?
R8	Oui. Nous encourageons la collaboration entre les institutions. Toutefois, seule une institution « principale » recevra des fonds. Les autres entités doivent être répertoriées

	comme des institutions partenaires et structurées administrativement comme des sous-traitants.
Q9	L'équipe peut-elle être composée de membres de différentes institutions ?
R9	Oui. Nous encourageons la collaboration entre les institutions.
Q10	Quels sont les critères d'évaluation pour décider des demandes que vous privilégieriez pour le financement ?
R10	Vous trouverez la liste complète des critères d'évaluation à la page 4 de l' <a href="#">appel à propositions</a> .
Q11	1. Comment définissez-vous les « grands projets transformationnels » ?  2. Avez-vous des exemples de ce que sont les petits/moyens projets par rapport aux grands projets ou projets transformationnels ?
R11	<p>Le Fonds et le groupe consultatif technique n'ont pas de définition stricte de « grand projet transformationnel ». Cependant, au sens large, les projets transformationnels doivent avoir un impact sur la disponibilité ou l'équité des données, permettant des progrès significatifs dans le domaine du TALN en Afrique.</p> <p>À titre d'exemple, un petit ensemble de données peut être orienté vers une seule tâche de TALN dans une seule langue, tandis qu'un ensemble de données moyen peut être plus approfondi ou complet dans une seule application ou une seule langue, ou encore un ensemble de données pour plusieurs tâches dans une seule langue. Un ensemble de données transformationnel peut traiter un problème commun de TALN ou permettre des avancées dans plusieurs langues ou plusieurs contextes. Un projet transformationnel pourrait également créer de nouveaux outils/plateformes d'annotation ou des processus communautaires permettant d'accélérer les efforts futurs dans ce domaine.</p> <p><b>Veillez noter que les projets « transformationnels » (dont le budget est compris entre 100 000 et 250 000 USD) doivent clairement démontrer au groupe consultatif technique leur impact potentiel sur le terrain, leur faisabilité et leur durabilité au-delà de la période de subvention.</b></p>
Q12	Une équipe peut-elle proposer d'aborder plusieurs ensembles de données ?
R12	Oui. Voir la question 1 pour des conseils sur les ensembles de données multiples. Dans un tel cas, l'équipe doit démontrer qu'elle dispose de la capacité nécessaire pour mener à bien plusieurs projets en même temps, si elle est sélectionnée.
Q13	Le « chercheur principal » doit-il posséder certaines qualifications spécifiques ?
R13	Non. Voir la question 4.
Q14	Le terme « institut de recherche » inclut-il les universités ?
R14	Oui. Vous trouverez la liste complète des critères d'éligibilité à la page 3 de l' <a href="#">appel à</a>

	<a href="#">propositions</a> .
Q15	Pourriez-vous expliquer ce que signifient les approbations nationales ?
R15	La réalisation de recherches dans certains pays peut nécessiter l'approbation des entités nationales compétentes (par exemple, d'un ministère de l'Éducation si la recherche est menée dans un établissement scolaire ou d'un ministère des Sciences et de la Technologie pour la recherche sur des sujets humains). Cette approbation n'est pas requise lors de la soumission de la proposition, toutefois, les candidats doivent démontrer qu'ils sont informés des processus d'approbation qu'ils devront suivre.
Q16	Y a-t-il une préférence ou une pondération pour les organisations basées en Afrique par rapport à une organisation hors d'Afrique ayant des partenariats « importants » en Afrique ? En d'autres termes, dans quelle mesure êtes-vous intéressé par le développement de cette compétence et de ces organisations sur le continent africain ?
R16	Il n'y a pas de préférence. Toutefois, le Lacuna Fund met l'accent sur une approche participative et les propositions seront évaluées sur la base de ce critère (voir les « Critères d'évaluation » dans l'appel à propositions).  Bien que nous accueillions favorablement les propositions émanant d'organisations basées en Afrique et d'organisations basées hors d'Afrique ayant des partenariats « importants » en Afrique, toutes les institutions doivent être engagées dans les aspects techniques des travaux proposés et partager les connaissances ou les avantages qui en découlent.
Q17	Bien qu'aucune application de bout en bout ne soit soutenue, est-ce que Lacuna soutient certains travaux de type prototype de vérification pour ces ensembles de données ? Par exemple, des modèles d'échantillons qui permettent à d'autres personnes d'utiliser facilement les ensembles de données.
R17	Tant que le but principal du développement de modèles est d'évaluer la qualité/utilité des ensembles de données développés (c'est-à-dire la validation), ou d'encourager davantage le travail sur les ensembles de données, une certaine budgétisation pour le développement de modèles est autorisée dans le cadre d'un projet soutenu.
Q18	Comment le Lacuna Fund envisage-t-il la « maintenance » des ensembles de données après la première année de financement ?
R18	Le Lacuna Fund n'héberge pas les ensembles de données, parce que nous voulons donner la flexibilité aux équipes d'héberger leurs données d'une manière qui a du sens pour ceux qui utiliseront les données. Les propositions peuvent inclure un poste budgétaire pour l'hébergement à perpétuité. En outre, toute idée des candidats visant à inciter une communauté à continuer à contribuer/maintenir un ensemble de données renforcerait une proposition. Voir les pages 5 et 11 de l' <a href="#">appel à propositions</a> pour plus d'informations sur la durabilité.
Q19	Encouragez-vous les collaborations avec des instituts de l'UE/des États-Unis travaillant dans des langues de l'Afrique subsaharienne ?

R19	Voir la question 16.
Q20	<p>1. Nous sommes une startup de [un pays d’Afrique du Nord] et nous sommes disposés à participer, sommes-nous éligibles ?</p> <p>2. Cet appel à propositions concerne-t-il toutes les langues africaines ou uniquement les langues d’Afrique subsaharienne ? Si c’est uniquement pour les langues d’Afrique subsaharienne, pouvons-nous participer en tant que [Nord-Africains] en créant une langue d’Afrique subsaharienne ?</p>
R20	<p>Vous trouverez la liste complète des critères d’éligibilité à la page 3 de l’<a href="#">appel à propositions</a>.</p> <p>La portée géographique de cet appel à propositions est limitée à l’Afrique subsaharienne, telle que définie par la Banque mondiale : <a href="https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups">https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups</a></p> <p>Toutefois, les organisations situées en dehors de l’Afrique subsaharienne peuvent s’associer à des organisations de la région pour des propositions.</p> <p>L’appel à propositions étant axé sur les langues d’Afrique subsaharienne, le contenu de votre proposition doit donc porter sur un ensemble de données relatives à ces langues.</p>
Q21	Avez-vous un avis particulier sur la langue haoussa en général ? Y a-t-il déjà eu des candidats ou des ensembles de données disponibles pour le haoussa ?
R21	La création d’un ensemble de données pour toute langue d’Afrique subsaharienne, y compris le haoussa, est prévue dans le cadre de l’appel à propositions. La responsabilité de l’étude et de la compréhension du paysage des ensembles de données existants pour une langue donnée incombe aux candidats. Veuillez consulter la page « Ressources » sur le site web du Fonds pour une liste non exhaustive des efforts existants.
Q22	Un enregistrement du webinaire du 2 octobre sera-t-il disponible ?
R22	Oui, un enregistrement du webinaire est disponible <a href="#">ici</a> .
Q23	Outre le financement publié concernant les ensembles de données pour les langues en Afrique subsaharienne, pourriez-vous nous faire savoir si le candidat peut présenter une nouvelle demande pour l’appel à propositions (soit en raison d’un changement de partenariat, soit pour une meilleure proposition) ?
R23	Les institutions peuvent faire plusieurs demandes pour des appels à propositions auprès du Lacuna Fund. Veuillez noter que cet appel d’offres récemment publié porte sur des ensembles de données relatifs à la langue, et non à l’agriculture.
Q24	1. Une proposition peut-elle porter sur plusieurs ensembles de données ? Est-il possible de traiter plusieurs langues à la fois dans un ensemble de données, par exemple un ensemble de données d’étiquetage morphosyntaxique pour plusieurs langues ?

	<p>2. Un candidat principal peut-il soumettre deux propositions, si une proposition porte sur un seul ensemble de données ?</p> <p>3. Quel est le matériel autorisé qui doit être inclus dans la proposition, y compris les machines ?</p>
R24	<p>1. Oui. Veuillez vous reporter à la question 1.</p> <p>2. Oui, un candidat principal peut présenter deux propositions. Dans un tel cas, l'équipe doit démontrer qu'elle dispose de la capacité nécessaire pour mener à bien les deux projets en même temps, si elle est sélectionnée.</p> <p>3. Les exigences relatives à la soumission d'un projet sont reprises dans l'appel à propositions à la section 4. La prise en compte des documents qui dépassent les éléments de la proposition spécifiés à la section 4 reste à la discrétion des évaluateurs.</p>
Q25	<p>Est-il possible d'avoir une seule proposition pour le même ensemble de données dans plusieurs langues ? Par exemple, puis-je soumettre une seule proposition si je crée un ensemble de données sur l'analyse des sentiments dans trois langues africaines ?</p>
R25	<p>Oui. Veuillez vous reporter à la question 1.</p>
Q26	<p>[Notre organisation internationale], par l'intermédiaire de sa [succursale en Afrique], a travaillé sur des outils de TALN pour les langues d'Afrique avec des établissements universitaires de la région au cours des six dernières années.</p> <p>Les subventions proposées par le Lacuna Fund nous permettront d'améliorer et de créer des ressources qui sont le résultat d'années de travail.</p> <p>Ma question est la suivante : [cette organisation internationale] est-elle éligible à une subvention dans le cadre de l'appel « langues » ?</p>
R26	<p>Oui, les organisations internationales peuvent soumettre un projet. Si leur siège n'est pas en Afrique, elles doivent avoir un partenariat conséquent avec une organisation dont c'est le cas.</p>
Q27	<p>Accepteriez-vous une candidature de la part d'un consortium d'organisations et de chercheurs ?</p> <p>Théoriquement, il y aurait différentes propositions d'ensembles de données au sein de la proposition du consortium, et bien qu'il y ait une organisation principale (avant tout pour tirer parti de sa capacité administrative), il pourrait y avoir plusieurs bénéficiaires secondaires de la subvention derrière cette institution principale qui serait idéalement affiliée à la proposition d'ensemble de données et serait donc l'organisation principale pour chaque proposition d'ensemble de données.</p> <p>Chaque proposition d'ensemble de données serait indépendante en termes de mise en œuvre et disposerait d'un budget distinct et clair pour ses activités, de sorte que si l'une</p>

	<p>d'entre elles était rejetée en raison de ses carences spécifiques, nous espérons que cela n'affecterait pas les autres propositions de la candidature du consortium.</p> <p>Grâce à une telle structure, une organisation disposant de capacités administratives en matière de subventions pourrait alors prêter ces capacités à d'autres organisations qui disposent des capacités techniques. Le consortium pourrait alors tirer parti de partenaires de recherche communs sur le plan juridique et éthique.</p>
R27	<p>Oui. Veuillez noter qu'une seule organisation « principale » recevra directement les fonds et que les partenaires doivent être structurés administrativement comme des sous-traitants. Veuillez vous reporter à la question 1 pour plus d'informations sur les propositions relatives à plusieurs ensembles de données connexes.</p> <p>Étant donné qu'un tel consortium pourrait constituer une proposition importante/transformationnelle (&gt;100 000 USD), veuillez vous reporter également à la question 11.</p> <p>En particulier, le Fonds accueille favorablement les propositions qui fournissent une justification convaincante à l'approche en consortium ou en partenariat plutôt que des propositions individuelles pour chaque élément du partenariat.</p>
Q28	<p>1. Pouvons-nous présenter plusieurs propositions ? Par exemple, présenter des demandes pour différentes langues en appliquant les mêmes méthodes dans le cadre de propositions distinctes ?</p> <p>2. Pouvons-nous présenter une demande avec des groupes non enregistrés en tant que partenaires ?</p>
R28	<p>1. Oui, voir la question 24. Voir la question 1 pour plus d'informations sur l'inclusion d'ensembles de données connexes dans une seule proposition.</p> <p>2. Le Lacuna Fund ne peut accorder de subventions qu'aux organisations principales éligibles. Les individus ou les organisations non enregistrées doivent s'engager dans le projet par l'intermédiaire d'un sponsor institutionnel.</p>
Q29	<p>Vous indiquez que les organisations éligibles doivent être des organismes à but non lucratif, des institutions de recherche et des entreprises sociales à but lucratif. Ma question est la suivante : qu'est-ce qui qualifie exactement une organisation en tant qu'entreprise sociale et, plus précisément, notre organisation est-elle éligible ? Nous sommes une entreprise, mais l'un de nos principaux objectifs est l'inclusion financière et, à cette fin, nous développons des modèles qui permettent aux personnes ayant un faible niveau d'anglais d'accéder aux technologies financières numériques dans leur propre langue maternelle. Nous voulons collecter et mettre à disposition des données qui aident à atteindre cet objectif. Pouvons-nous proposer notre candidature ?</p>
R29	<p>Il n'existe pas de désignation officielle pour une entreprise sociale ; toute organisation à but lucratif dont la mission est liée au bien social est invitée à proposer sa candidature.</p>

	<p>Veillez consulter les autres critères d'éligibilité à la page 3 de l'<a href="#">appel à propositions</a> afin de vous assurer que vous répondez à tous les autres critères.</p>
<p>Q30</p>	<ol style="list-style-type: none"> <li>1. La proposition peut-elle porter sur plus d'une langue ?</li> <li>2. Pouvons-nous soumettre une proposition qui aborde deux langues ?</li> <li>3. Pour les corpus parallèles, pouvons-nous simplement utiliser l'anglais comme langue principale, les données étant donc par exemple en anglais-[langue africaine] ?</li> <li>4. Pouvons-nous travailler sur un ensemble de données non symétrique, mais dans le même domaine pour l'apprentissage non supervisé ?</li> <li>5. Le salaire du personnel peut-il être mensuel au lieu d'être calculé par heure / par jour ?</li> <li>6. Tous les salaires doivent-ils être TTC, par exemple dans le cas où il faudrait engager un consultant ?</li> <li>7. Pourriez-vous fournir un exemple de modèle de budget rempli ? Il servira de guide pour remplir le budget correctement.</li> <li>8. Si le plan de collecte des données prévoit d'inciter les gens à partager des données et de les récompenser en leur accordant un crédit téléphonique ou autre, ces coûts doivent-ils être indiqués ?</li> <li>9. La location de bureaux ou d'espaces communs pour effectuer le travail peut-elle être considérée comme un coût indirect ?</li> <li>10. Quels sont les différents budgets (max. et min.) selon les différentes portées de projet ?</li> <li>11. Disposons-nous au moins d'un ensemble de données de la taille minimale requise par langue au cas où l'on créerait de nouvelles données ? Par exemple, le nombre de phrases ou de mots, etc.</li> <li>12. Pouvons-nous utiliser GitHub pour le stockage des données après la fin du projet ? Un entrepôt de données gratuit ?</li> </ol>
<p>R30</p>	<ol style="list-style-type: none"> <li>1. Voir la question 1.</li> <li>2. Voir la question 1.</li> <li>3. Les ensembles de données proposés doivent être adaptés aux cas d'utilisation indiqués dans la proposition.</li> </ol>

	<p>4. Oui. Voir la section 3 « Objet et besoins » de l’appel à propositions, notamment : « les recueils de textes étiquetés et non étiquetés pour les utiliser comme données de formation ».</p> <p>5. Oui, c’est possible. Veuillez indiquer les unités dans le budget.</p> <p>6. Oui, ils doivent être TTC.</p> <p>7. Un exemple de budget sera disponible sur la page de candidature ici.</p> <p>8. Dans le budget, ce coût sera inclus dans les autres coûts.</p> <p>9. Oui, le loyer et les espaces communs sont considérés comme des coûts indirects et relèveraient de l’allocation de 12 % pour les coûts indirects.</p> <p>10. Veuillez vous reporter à la section « Budget » de l’appel à propositions, à la page 12 : « Nous prévoyons des budgets dans une fourchette de 20 000 à 100 000 USD pour les projets de petite à moyenne taille, et des budgets pouvant atteindre 500 000 USD pour les projets de grande ampleur et complexes. Il convient de noter que le Groupe consultatif technique évaluera la faisabilité et la pertinence du budget ainsi que sa cohérence avec l’exposé de la demande de subvention dans le cadre des critères de sélection. »</p> <p>11. Voir la question 2.</p> <p>12. Les propositions doivent comprendre un plan de durabilité, y compris une solution d’hébergement qui remplit les critères d’« accessibilité » et de « durabilité » inclus dans la section « Critères d’évaluation » de l’appel à propositions.</p>
Q31	<p>1. Quelle est la position du Fonds sur l’externalisation de certains aspects des activités du projet à un prestataire de services compétent dont l’implication permettra d’accélérer le projet ?</p> <p>2. Quel est le plafond budgétaire pour les services de consultance ?</p>
R31	<p>1. Le Fonds n’a pas de position. Nous reconnaissons que les propositions pourraient être renforcées par la participation de consultants compétents (par exemple, des linguistes ou des spécialistes juridiques). Toutefois, les candidats devraient examiner les principes du Lacuna Fund, notamment l’accent mis sur une approche participative et les bonnes pratiques en matière d’éthique, y compris les normes de travail équitables. Les structures de la proposition et du partenariat doivent refléter ces principes.</p> <p>2. Nous n’avons pas de limite budgétaire pour les consultants.</p>
Q32	<p>1. Combien de projets sont susceptibles d’être financés dans le cadre du LACUNA FUND : Appel à propositions 2020 pour les langues ?</p> <p>2. Est-il acceptable qu’une collection d’ensembles de données comprenne des données de formation pour la reconnaissance vocale et l’automatisation de l’analyse morphologique ?</p>

	<p>En d’autres termes, est-il permis de produire une collection hybride d’ensembles de données ?</p> <p>3. Qui est responsable de la gestion financière de la subvention ? Est-ce le chercheur principal ou le département financier de l’institution d’accueil ?</p>
R32	<p>1. Voir la question 3.</p> <p>2. Oui. Voir la question 1.</p> <p>3. Cette détermination est de la responsabilité du candidat. Le bénéficiaire de la subvention est responsable de la gestion financière et de la conservation des dossiers (s’ils sont financés).</p>
Q33	<p>Toutes les personnes mentionnées dans une demande conjointe doivent-elles être affiliées à une « entité à but non lucratif, un institut de recherche ou une entreprise sociale à but lucratif » ou seulement le demandeur principal ? Existe-t-il une définition fixe de ce qu’est une « entreprise sociale à but lucratif » ?</p>
R33	<p>Voir les questions 28 et 29.</p>