

# Baybars Hackathon – Call for applications

French Institute of the Near-East (Ifpo)

Beirut

September 26-29 2023

En français ci-dessous

بالعربية أدناه

As part of the research program “Popular Literatures of the Levant » (*Littératures populaires du Levant – LiPoL*) and in partnership with the National Library of France (BnF), the startup company Calfa and the media library of the Mediterranean House of Human Sciences (MMSH), the French Institute of the Near-East (Ifpo) is organizing from September 2023 in Beirut a *hackathon* to transcribe a corpus of manuscripts from the *Sīrat Baybars* in its Damascene recension. This corpus consists of 542 storytellers' notebooks corresponding to around 60,000 pages of text in Middle Arabic. A third of the corpus has already been transcribed for publication by the Presses de l'Ifpo.

This "Baybars Hackathon" aims to obtain an HTR (Handwritten Text recognition) model for the automatic transcription of the entire corpus, thanks to the initial manual transcription of a sample of the manuscripts.

The project will involve a team of 12 people selected through an application process from Master's and PhD students enrolled in a Lebanese university and specialized in Arabic language and literature, History or codicology. Students from French universities who are already members of the LiPoL program will also attend the hackathon.

\*\*\*

A preliminary four-day training will take place at Ifpo Beirut from September 26 to September 29, 2023. This training will address a range of issues relating to Arabic-language manuscripts in general, and to manuscripts of popular literature in particular, and to the digital humanities, around four themes:  
- HTR training provided by Calfa, a company specialized in OCR and HTR for non-Latin languages, which will support us throughout the project;

- an introduction to Arabic manuscripts through codicology and palaeography;
- a training in Middle Arabic, a register of Arabic language situated between standard and dialect, widely used by storytellers of popular and middle Arabic literature, which the *sīra* is part of;
- an introduction to popular and middle Levantine literature, to the handwritten sources of this literature, to the *sīra shaʿbiyya* genre and *Sīrat Baybars*.

At the end of the training days, participants will leave with a set of pages to transcribe on their own during the following weeks in order to train the Artificial Intelligence which will generate the HTR model.

\*\*\*

The “Baybars Hackathon” is part of a range of numerous scientific initiatives which aim to improve the place of Arabic language in the Digital Humanities field and in the scientific research. It follows the "Alexander Hackathon", organized between November 2022 and May 2023 around manuscripts of the Alexander Romance in Arabic. The Alexander Hackathon produced textual data which will serve the critical edition of the romance. The event involved students from Sorbonne University, Strasbourg University, Lorraine University and Aix-Marseille University (see: <https://lipol.hypotheses.org/790>).

## Partnership

The “Baybars Hackathon” is organized as part of the research program “Popular Literatures of the Levant » (*Littératures populaires du Levant – LiPoL*), which is funded by the French National Research Agency (ANR), in partnership with the National Library of France (BnF) and the Mediterranean House of Human Sciences (MMSH).

## Technical and scientific supervision

Taos BABOUR, Research Engineer, head of the Digital Humanities department at the Ifpo, Beirut: [www.ifporient.org/taos-babour](http://www.ifporient.org/taos-babour); email: [t.babour@ifporient.org](mailto:t.babour@ifporient.org)

Claire CIALONE-GREGOIRE, Digital Humanities Engineer for LiPoL program: [cv.hal.science/claire-cialone-gregoire](http://cv.hal.science/claire-cialone-gregoire); email: [claire.cialone-gregoire@cnrs.fr](mailto:claire.cialone-gregoire@cnrs.fr)

Iyas HASSAN, University Professor in Arabic Literature at Sorbonne University and member of the Ifpo, head of LiPoL program: [www.orient-mediterranee.com/member/hassan-iyas](http://www.orient-mediterranee.com/member/hassan-iyas); email: [iyas.hassan@cnrs.fr](mailto:iyas.hassan@cnrs.fr)

## **Application and validation of the participation**

The Baybars Hackathon is open to Master's and PhD students enrolled in Lebanese universities in Arabic language and literature, in History, codicology, or any other subject related to Arabic studies and Arabic handwritten sources.

The workshop is open to beginners and no previous experience with manuscripts, HTR, popular literature or Middle Arabic is required. However, perfect knowledge of written standard Arabic and its grammatical rules is mandatory to participate in the hackathon and benefit from the training.

The applicant shall provide a recommendation letter from his/her course supervisor, explaining how the participation to the hackathon could benefit his/her university education.

At the end of the hackathon in 2024, participants will receive a certificate of attendance from the Ifpo. With the prior agreement of each university's department, participation in the hackathon could be validated as part of the student's course.

A commitment from the participants will be required for the entire project, from September 2023 to February 2024 (see provisional schedule below), in order to regularly check the quality of the transcription produced by the HTR.

Shared room accommodation at the Ifpo may be available for participants from universities outside Beirut (limited places).

## **Provisional schedule**

**September 8, 2023:** Application deadline

**September 26-29, 2023:** Initial training at Ifpo Beirut

**September 29 – October 31, 2023:** Transcription work (remotely)

**October 31, 2023:** Submission of all samples

**November 7, 2023:** Verified transcripts will be sent to Calfa for processing

**February 1, 2024:** Reception of the text transcribed by the HTR model developed by Calfa

## **In order to apply to the Baybars Hackathon**

Please send your resume with the recommendation of your course supervisor to :

Taos BABOUR – [t.babour@ifporient.org](mailto:t.babour@ifporient.org)

Claire CIALONE-GREGOIRE – [claire.cialone-gregoire@cnrs.fr](mailto:claire.cialone-gregoire@cnrs.fr)

Iyas HASSAN – [iyas.hassan@cnrs.fr](mailto:iyas.hassan@cnrs.fr)

# بَرْمَجَان بِيِيرِص

## Hackathon Baybars

المعهد الفرنسي للشرق الأدنى

بيروت

٢٦، ٢٧، ٢٨، ٢٩ أيلول ٢٠٢٣

En français ci-dessous

ضمن أنشطة برنامج "آداب المشرق الشعبية" (LiPoL)، ينظم المعهد الفرنسي للشرق الأدنى (إفبو) بالتعاون مع المكتبة الوطنية الفرنسية ومؤسسة Calfa والمكتبة السمعية والبصرية في بيت المتوسّط للعلوم الإنسانيّة (MMSH)، اعتباراً من أيلول ٢٠٢٣ في بيروت، بَرْمَجَانًا لرقن محتوى مخطوطات سيرة الملك الظاهر بيبيرص حسب الرواية الشامية بصورة آلية. تتكون هذه المخطوطات من مجموعة دفاتر حكايتين دمشقيين مكتوبة بلغة عربيّة وسطى، يبلغ عددها ٥٤٢ دفترًا، تشكّل حوالي ٦٠٠٠٠ صفحة مخطوطة. ولقد نُسخ حوالي ثلث هذه الدفاتر يدويًا أثناء عمليات تحقيق السيرة وإصدارها بطبعة علميّة عن منشورات الإفبو بين عامي ٢٠٢٢-٢٠٠٠.

يهدف "برمجان بيبيرص" إلى تطوير نموذج رقمي للتعرف على الخط المكتوب يدويًا ( *Handwritten Text* recognition, HTR) سيستخدم لنسخ كامل مخطوطات السيرة بصورة آليّة، وذلك انطلاقًا من عيّنة تنسخ وتدقّق يدويًا.

يتكوّن فريق "برمجان بيبيرص" من اثني عشر راقنًا وراقنة سيختارون/يخترن من طلاب وطالبات الحلقتين الجامعيّتين الثانية (ماستر) والثالثة (دكتوراه) في لبنان، من المختصّين والمختصات باللغة العربية وآدابها، بالتاريخ أو بعلم المخطوطات، وسينضمّ إليهم/إليهنّ مجموعة من طلاب وطالبات الدراسات العليا في فرنسا من أعضاء مشروع "آداب المشرق الشعبيّة" (LiPoL).

\*\*\*

تنطلق أعمال "برمجان بيبيرص" عبر أربعة أيام من التدريب، ٢٦ و ٢٧ و ٢٨ و ٢٩ أيلول/سبتمبر في بيروت، يتعرّف المشاركون والمشاركات خلالها إلى إشكاليّات متعلّقة بالمخطوطات العربيّة بشكل عام، بمخطوطات الأدب الشعبي بشكل خاص وبالإنسانيّات الرقمية، وذلك عبر أربعة محاور :

- اكتشاف تقنيات التعرف على الخط المكتوب يدوياً (*Handwritten Text recognition, HTR*) وتدريب أولي على استخدام البرنامج تشرف عليه مؤسسة Calfa الفرنسية الرائدة في مجال التعرف الآلي على الخط المكتوب (HTR) والمطبوع (OCR) بلغات غير لاتينية، والتي سترافق أعمال البرمجان حتى نهايته.
  - مدخل إلى المخطوطات العربية وإلى علم المخطوطات.
  - مدخل إلى "العربية الوسطى"، المقام اللغوي واسع الانتشار في مخطوطات الأدب الشعبي والأوسط.
  - مدخل إلى الآداب الشعبية والوسطى العربية، إلى مصادرها المكتوبة، إلى السيرة الشعبية كنوع أدبي وإلى سيرة الملك الظاهر.
- بعد نهاية أيام التدريب، سيُكف كل من أعضاء الفريق بنسخ مجموعة من الصفحات بشكل مستقل خلال الأسابيع التالية وإدراجها في برنامج التعرف الآلي.

\*\*\*

يندرج "برمجان بيبيرص" ضمن مجموعة من الفعاليات العلميّة الهادفة إلى تحسين موقع اللغة العربية بمختلف مقاماتها في الفضاء الافتراضي، وإلى تطوير الأدوات الرقمية القادرة على تسهيل البحث العلمي في حقل الدراسات العربية وفتح آفاق جديدة له. ويرتبط "برمجان بيبيرص" مباشرة بالأعمال التي أنجزت في "برمجان الإسكندر" (Hackathon Alexandre) الذي عملت فيه مجموعة من الطلاب والطالبات في جامعات السوربون، ستراسبورغ، إكس-مرسيليا واللورين، بين تشرين الثاني ٢٠٢٢ وأيار ٢٠٢٣ استناداً إلى مخطوطات سيرة الإسكندر الشعبية العربيّة، والتي ستبدأ أعمال تحقيقها العلمي في نهاية العام ٢٠٢٣ (لمزيد من المعلومات عن "برمجان الإسكندر": <https://lipol.hypotheses.org/790>)

\*\*\*

## المنظّمون

ينظّم "برمجان بيبيرص" المعهد الفرنسي للشرق الأدنى ضمن أنشطة برنامج أبحاث "آداب المشرق الشعبيّة" (LiPoL) الذي تموّله الوكالة الوطنية للبحث العلمي بفرنسا، بالاشتراك مع المكتبة الوطنية الفرنسيّة وبيت المتوسّط للعلوم الإنسانية.

## الإشراف التقني والعلمي

- طاوس بابور، مهندسة أبحاث، مسؤولة قسم الإنسانيّات الرقمية في المعهد الفرنسي للشرق الأدنى، بيروت : [www.ifporient.org/taos-babour](http://www.ifporient.org/taos-babour) ؛ البريد الإلكتروني : [t.babour@ifporient.org](mailto:t.babour@ifporient.org)
- كلير تشالون - غريغوار، مهندسة دراسات، مسؤولة الإنسانيّات الرقمية في مشروع آداب المشرق الشعبيّة : [cv.hal.science/claire-cialone-gregoire](http://cv.hal.science/claire-cialone-gregoire) ؛ البريد الإلكتروني : [claire.cialone-gregoire@cnrs.fr](mailto:claire.cialone-gregoire@cnrs.fr)
- إياس محسن حسن، أستاذ الأدب العربي بجامعة السوربون/المعهد الفرنسي للشرق الأدنى، مسؤول برنامج "آداب المشرق الشعبيّة" : [www.orient-mediterranee.com/member/hassan-iyas/](http://www.orient-mediterranee.com/member/hassan-iyas/) ؛ البريد الإلكتروني : [iyas.hassan@cnrs.fr](mailto:iyas.hassan@cnrs.fr)

## التسجيل

"برمجان بيبيرص" مفتوح لطلاب وطالبات الحلقتين الثانية (ماستر) والثالثة (دكتوراه)، المسجّلين/ات في الجامعات اللبنانية، في اختصاص اللغة العربية وآدابها، التاريخ، علم المخطوطات أو أي من الحقول الأخرى المرتبطة بالدراسات العربيّة وبدراسة المصادر المخطوطة العربيّة.

لا تتطلب المشاركة بالبرمجان معرفة مسبقة بالمخطوطات العربية ولا بتقنيات التعرف الآلي على الخط المكتوب، ولا بإشكاليات العربية الوسطى أو الآداب الشعبيّة، مع أن الإلمام بهذه العناوين محبّب. لكنّ امتلاك اللغة العربية المكتوبة بشكل كامل، ومعرفة قواعد الفصحى معرفة صارمة شرطان أساسيان للاشتراك في هذا العمل وللاستفادة منه.

يُشترط للترشح أن يحتوي الملفّ على رسالة من الأستاذة/المشرفة/ة توضح فائدة الاشتراك في هذا المشروع من وجهة نظر الحقل العلمي الذي يتخصص فيه المشارك/ة وما ستضيفه هذه التجربة إلى تكوينه/ها

إدراج هذا العمل رسمياً في برنامج تكوين الطلاب والطالبات عائد حصراً إلى جامعاتهم/هنّ، علماً بأن المعهد الفرنسي للشرق الأدنى سيقدّم، بعد انتهاء الأعمال في ٢٠٢٤، وثيقة تصف المهام التي أوكلت إلى كل مشارك/ة والأعمال التي أنجزها/تها، وذلك لتسهيل الحصول على معادلة أو إدراج المهارات المكتسبة في أحد المقرّرات، إن ارتأت الجامعة المعنيّة ذلك.

الاشتراك ببرنامجان ببيبرص يعني الالتزام به حتى نهايته في شباط/فبراير ٢٠٢٤ (انظر/ي الجدول الزمني أدناه)، إذ سيطلب من المشاركين/ات، بعد الفراغ من نسخ الصفحات الموكلة إليهم/هنّ، أن يقوموا/يقمن بشكل دوري بأعمال تدقيق للنصوص التي سينتجها الراقن الآلي.

سيمكن للمشاركين/ات من جامعات خارج بيروت المبيت في غرفة مشتركة في المعهد الفرنسي للشرق الأدنى (الأماكن محدودة).

## الجدول الزمني

٨ أيلول/سبتمبر الساعة السادسة مساءً بتوقيت بيروت : آخر موعد لقبول الترشيحات.

٢٦-٢٩ أيلول/سبتمبر : إطلاق أعمال "برنامج ببيبرص" في بيروت (سيعلن البرنامج العلمي قريباً).

٢٩ أيلول/سبتمبر إلى ٣١ تشرين الأول/أكتوبر : استكمال أعمال الرقن اليدوي عن بعد.

٣١ تشرين الأول/أكتوبر : آخر موعد لتسليم العينات المنسوخة.

٧ تشرين الثاني/نوفمبر : إرسال جميع العينات، بعد تدقيقها، إلى مؤسسة Calfa للبدأ بتطوير النموذج الرقمي.

١ شباط/فبراير ٢٠٢٤ : استلام النصوص المرقونة آلياً.

## للاشتراك في برنامج ببيبرص

إرسال السيرة الذاتية مرفقة برسالة من الأستاذة/المشرفة إلى العناوين التالية :

طاولس بابور / Taos BABOUR – [t.babour@ifporient.org](mailto:t.babour@ifporient.org)

كلير تشالون – Claire CIALONE-GREGOIRE – [claire.cialone-gregoire@cnrs.fr](mailto:claire.cialone-gregoire@cnrs.fr)

إياس محسن حسن / Iyas HASSAN – [iyas.hassan@cnrs.fr](mailto:iyas.hassan@cnrs.fr)

# Hackathon Baybars - Appel à candidatures

Institut français du Proche-Orient  
Beyrouth

26, 27, 28, 29 septembre 2023

Dans le cadre du programme ANR « Littératures populaires du Levant » (LiPoL), en partenariat avec la Bibliothèque nationale de France (BnF), l'entreprise Calfa et la Médiathèque SHS de la Maison Méditerranéenne des sciences de l'Homme (MMSH), l'Institut français du Proche-Orient (Ifpo) organisera à partir de septembre 2023 à Beyrouth un hackathon pour la transcription d'un corpus de manuscrits de la *sīra* de Baybars dans sa recension damascène. Ce corpus est constitué d'un ensemble de 542 cahiers de conteurs correspondant à environ 60.000 pages de texte en moyen-arabe, dont un tiers a déjà été transcrit pour publication aux Presses de l'Ifpo.

Ce « Hackathon Baybars » visera à obtenir un modèle HTR (*Handwritten Text recognition*) pour la transcription automatique de l'ensemble du corpus, grâce à la transcription manuelle préalable d'un échantillon des manuscrits.

Le projet mobilisera une équipe de 12 personnes sélectionnées après candidature parmi les étudiant·e·s de Master et de Doctorat spécialisé·e·s en langue et littérature arabes, en histoire ou en codicologie, inscrit·e·s dans une université libanaise. Participeront également au hackathon les étudiant·e·s d'universités françaises déjà membres du programme ANR LiPoL.

\*\*\*

Le hackathon sera amorcé par quatre journées de formation du 26 au 29 septembre 2023 à Beyrouth. Ces journées seront l'occasion d'aborder différentes problématiques liées aux manuscrits en langue arabe en général, aux manuscrits de littérature populaires en particulier, et aux humanités numériques, et ce à travers quatre axes :

- une formation à l'HTR par Calfa, entreprise spécialisée dans l'OCR et l'HTR des langues non-latines et qui nous accompagnera tout au long du projet ;
- une introduction au manuscrit arabe à travers la codicologie et la paléographie ;
- une formation au moyen-arabe, registre de l'arabe situé entre le standard et le dialectal et employé largement par les conteurs de littérature populaire et médiane arabe dont la *sīra* ;



- une introduction à la littérature populaire et médiane levantine, aux sources écrites de cette littérature, au genre *sīra shaʿbiyya* et à la *sīra* de Baybars.

A l'issue des journées de formation, les participant·e·s repartiront avec un ensemble de pages à transcrire en autonomie durant les semaines qui suivront afin d'entraîner l'intelligence artificielle qui générera le modèle HTR.

\*\*\*

Le *Hackathon Baybars* fait partie des nombreuses initiatives scientifique visant à améliorer la place de la langue arabe dans les humanités numériques dans la recherche scientifique. Il se situe dans la lignée de ce qui avait été réalisé lors du *Hackathon Alexandre* qui a eu lieu entre novembre 2022 et mai 2023 autour de manuscrits du *Roman d'Alexandre* en langue arabe. Le *Hackathon Alexandre* a produit des données textuelles qui seront mobilisées dans les travaux d'édition critique de cette œuvre. Il avait réuni, à la MMSH à Aix-en-Provence, une équipe d'étudiant·e·s de Sorbonne Université, Université de Strasbourg, Université de Lorraine et Aix-Marseille Université (voir : <https://lipol.hypotheses.org/790>).

## **Organisateurs et partenaires**

Le *Hackathon Baybars* est organisé par l'Institut français du Proche-Orient (Ifpo) dans le cadre des activités scientifiques du programme « Littératures populaires du Levant » (LiPoL) financé par l'Agence Nationale de la Recherche (ANR), en partenariat avec la Bibliothèque nationale de France (BnF).

## **Supervision technique et scientifique**

Taos BABOUR, ingénieure de recherche, responsable du département des humanités numériques à l'Ifpo, Beyrouth : [www.ifporient.org/taos-babour](http://www.ifporient.org/taos-babour) ; email : [t.babour@ifporient.org](mailto:t.babour@ifporient.org)

Claire CIALONE-GREGOIRE, ingénieure d'études en humanités numériques pour le programme ANR LiPoL : [cv.hal.science/claire-cialone-gregoire](http://cv.hal.science/claire-cialone-gregoire) ; email : [claire.cialone-gregoire@cnrs.fr](mailto:claire.cialone-gregoire@cnrs.fr)

Iyas HASSAN, professeur en littérature arabe médiévale à Sorbonne Université, membre de l'Ifpo, responsable du programme ANR LiPoL : [www.orient-mediterranee.com/member/hassan-iyas](http://www.orient-mediterranee.com/member/hassan-iyas) ; email : [iyas.hassan@cnrs.fr](mailto:iyas.hassan@cnrs.fr)

## Inscription et validation de la formation

Le *Hackathon Baybars* est destiné aux étudiant·e·s inscrit·e·s en Master et en Doctorat des universités libanaises et suivant un cursus en langue et littérature arabes, en histoire, en codicologie ou tout autre discipline en lien avec les études arabes et les sources manuscrites arabes.

Aucune expérience antérieure en codicologie, dans le traitement numérique des manuscrits, en littérature populaire ou en arabe médian n'est requise. La parfaite maîtrise de l'arabe écrit et de la grammaire de l'arabe standard est toutefois indispensable pour participer au hackathon et pour tirer profit de la formation.

Les candidat·e·s doivent faire état d'une recommandation écrite de leur responsable de formation ou directeur/directrice de recherche, expliquant l'intérêt de l'étudiant·e à participer à ce projet du point de vue de la formation qu'il/elle suit dans son université.

A l'issue du hackathon en 2024, les participant·e·s recevront une attestation de participation de la part de l'Ifpo. En concertation avec les responsables de formation de chaque université, la participation au hackathon pourrait donner lieu à la validation d'un enseignement équivalent.

Un engagement de la part des participant·e·s sera demandé sur l'ensemble du projet, de septembre 2023 à février 2024 (voir le calendrier provisoire ci-dessous), afin de procéder périodiquement au contrôle de la qualité de la transcription issue de l'HTR.

Une solution d'hébergement en chambre partagée à l'Ifpo pourra être proposée aux participant·e·s venant d'universités hors de Beyrouth (places limitées).

## Calendrier (sous réserve de modifications)

**8 septembre 2023, 18h heure Beyrouth** : date limite de candidature

**26-29 septembre 2023** : Début des travaux du *Hackathon Baybars* ; journées de formation à Beyrouth (programme à venir).

**29 sept. - 31 octobre 2023** : poursuite de la transcription à distance

**31 octobre 2023** : rendu de l'ensemble des échantillons

**7 novembre 2023** : envoi des transcriptions vérifiées à Calfa pour le traitement

**1<sup>er</sup> février 2024** : réception du texte transcrit par le modèle HTR élaboré par Calfa

## **Pour participer au hackathon Baybars :**

Envoyer son CV accompagné de la lettre de recommandation de l'enseignant·e référent·e aux adresses suivantes :

Taos BABOUR – [t.babour@iforient.org](mailto:t.babour@iforient.org)

Claire CIALONE-GREGOIRE – [claire.cialone-gregoire@cnrs.fr](mailto:claire.cialone-gregoire@cnrs.fr)

Iyas HASSAN – [iyas.hassan@cnrs.fr](mailto:iyas.hassan@cnrs.fr)

