

PRÉSENTATION DU PROJET

Organisé par le Campus Condorcet et le médialab de Sciences Po, cet événement s'inscrit dans le développement de l'offre de service de l'Humathèque, autour des données du web.

Le temps d'un *sprint*, réunis en quatre équipes et autant de thématiques, chercheurs, ingénieurs, bibliothécaires collaborent afin de produire des corpus web en sciences humaines et sociales.

Vous souhaitez découvrir comment se déroule un *data sprint*, les domaines explorés, voire tester le processus de constitution d'un corpus web ? Plusieurs modalités de participation vous sont proposées.

Élaborez votre programme à la carte, selon vos disponibilités et votre curiosité.

PROGRAMME

MARDI 18 NOVEMBRE

10h00 - 12h00 → Présentation du *crawler Hyphe*

15h30 - 17h00 → Expérimenter la constitution d'un corpus

MERCREDI 19 NOVEMBRE

11h00 - 11h30 → Visite guidée

13h30 - 14h00 → Visite guidée

14h00 - 15h30 → Expérimenter la constitution d'un corpus

JEUDI 20 NOVEMBRE

10h30 - 12h00 → Expérimenter la constitution d'un corpus

15h30 - 17h00 → Restitution

17h00 - 18h00 → Conférence de clôture

18h00 - 19h00 → Cocktail

DÉTAIL DES ACTIVITÉS

PRÉSENTATION DU CRAWLER HYPHE PAR BENJAMIN OOGHE-TABANOU (2H, PLÉNIÈRE, 20 PERS. MAX.)

Découvrez *Hyphe*, un outil de création et nettoyage de corpus web reposant sur un *crawler* orienté pour la recherche développé par le médialab (Sciences Po).

VISITE GUIDÉE (30MIN., EN GROUPE, 10 PERS. MAX. PAR CRÉNEAU)

À l'issue de cette visite, vous comprendrez ce qu'est un *data sprint*, pourquoi l'Humathèque mène cette expérimentation et quelles sont les différentes étapes pour constituer un corpus web. Nous discuterons également des sujets choisis.

EXPÉRIMENTER LA CONSTITUTION D'UN CORPUS (1H30, EN GROUPE, 4 PERS. MAX PAR CRÉNEAU)

Mettez les mains dans le cambouis. Sélectionnez et indexez des données pour visualiser un corpus en train de se construire.

RESTITUTION (1H30, PLÉNIÈRE, 20 PERS. MAX.)

Après 2 jours de travail, où en sont les équipes et à quoi ressemblent les corpus ? Quels sont les écueils rencontrés, les questions qu'elles se sont posées et les premières hypothèses émises ?

ARCHIVAGE DES DONNÉES DU WEB, QUELS ENJEUX POUR LA RECHERCHE EN SHS ? SOPHIE GEBEIL

CONFÉRENCE DE CLÔTURE (1H00, PLÉNIÈRE, 25 PERS. MAX.)

Sophie Gebeil est maître de conférences au laboratoire TELEMMe (AMU, CNRS) en histoire contemporaine. Ses recherches portent sur l'étude des pratiques mémorielles fondées sur l'utilisation des archives du web comme source pour l'historien. Elle coordonne l'atelier *Visual Studies et humanités numériques en Méditerranée* et est lauréate d'une chaire Médiation scientifique à l'Institut Universitaire de France.

La conférence sera suivie d'un cocktail.