Plateformes

Présentation des plateformes

La Maison des sciences de l'homme en Bretagne dispose de trois plateformes :

- la plateforme **Humanités numériques** et la **plateforme universitaire de données** mutualisent un ensemble de services et d'infrastructures qui vise à **accompagner les chercheurs dans la gestion de leurs données qualitatives et quantitatives**.
- La **Plateforme d'Analyse du Comportement et des Interactions de l'Humain dans son Milieu (PACIHM)** accompagne la communauté scientifiques en SHS dans la mise en place d'études autour du comportement humain.

Trois plateformes technologiques

La plateforme **Humanités numériques** apporte un **soutien dans le traitement, l'analyse et la conservation des données de recherche** (textuelles, iconographiques, audiovisuelles...). Elle est équipée d'un matériel de numérisation adapté à la production de corpus. Cette plateforme est la **porte d'entrée régionale vers les outils et services développés par l'Infrastructure de recherche (IR*) Huma-Num**. Les deux volets d'expertise développés par la plateforme Humanités numériques sont :

1. Traitement, analyse et édition de corpus textuels
2. Valorisation des archives de la recherche en SHS

La **plateforme universitaire de données - Bretagne** (PUD-B) s'inscrit dans le cadre de l'infrastructure de recherche (IR*) PROGEDO dédiée à la **production** et à la **gestion** des **données quantitatives en sciences sociales**. Elle **favorise l'accès et l'utilisation des données issues de la statistique** publique française, européenne et internationale (grandes enquêtes, bases de données internationales, centre d'accès sécurisé aux données...). Deux volets d'expertise sont développés par la plateforme :

1. recherche et production de données quantitatives
2. traitements et analyses statistiques liées aux données quantitatives

La **Plateforme d'Analyse du Comportement et des Interactions de l'Humain dans son Milieu (PACIHM)** accompagne la communauté scientifiques en SHS dans la mise en place d'études autour du comportement humain.

Pour répondre à ces objectifs, PACIHM met à disposition **des moyens humains** (temps ingénieurs) et **techniques** (matériel, locaux, panel testeurs, etc.).

Services des plateformes

Voici une synthèse des principaux services et outils proposés par les plateformes technologiques de la MSHB. Pour plus de détails, consultez les onglets dédiés.

Plateforme Humanités numériques (HN)

- Conseil sur les méthodes et outil numériques pour la conception d'un projet en Humanités numériques
- Accompagnement de projets en SHS à toutes les étapes du cycle de vie des données (collecte, traitement, stockage, sauvegarde, diffusion, valorisation)
- Soutien technique en édition numérique et en traitement automatique des langues

Production et analyse de corpus textuels :

- Numérisation de corpus (mise à disposition d'équipements de numérisation)
- Prise en main de logiciels de reconnaissance automatique de textes (ATR)

  : eScriptorium, Arkindex

- Prise en main TACTEO : plateforme de transcription collaborative
- Encodage XML-TEI
- Prise en main de logiciels d'analyse textuelle : Iramuteq, TXM, NVivo, QualCoder
- Traitements automatisés avec Python (webscraping, alignement de vocabulaires)
- Tâches TAL ou IA avec NLTK, spaCy, Transformers (tokenisation, fréquence de mots, classification, reconnaissance d'entités nommées, analyse de sentiments, question-réponse sur corpus, etc.)

Traitement et éditorialisation des archives pour la recherche en SHS

- constitution de corpus textuels et/ou iconographiques
- conseils sur la structuration de bases de données qualitatives
- documentation, dépôt dans un entrepôt, signalement et éditorialisation de corpus qualitatifs avec les outils de l'IR* Huma-Num

Relais de l'IR* Huma-Num en Bretagne :

- Information sur la grille de services Huma-Num
- Transmission à Huma-Num des besoins locaux

Accès et aide à l'utilisation des services mis en œuvre par Huma-Num (notamment pour les services suivants) :

- **Isidore** – assistant de recherche en SHS
- **Sharedocs** – stockage et partage de données pendant la recherche, avec des outils de traitement intégrés (OCR, Whisper,transcodage audio/ vidéo)

- **Nakala** – entrepôt de données de référence pour les SHS et son module d'éditorialisation **Nakala_Press**
- **Cocoon** (+ le Laboratoire de Langues et civilisations à tradition orale (LACITO) et le Laboratoire Ligérien de Linguistique (LLL)) – entrepôt de données pour les corpus oraux
- **Heurist** –service de bases de données relationnelles
- **TXM** –service d'analyse de corpus de textes
- **Gitlab** – hébergement, versionning et partage de code
- **Open Theso** –logiciel libre de gestion de thésaurus multilingue
- Informations sur l'accès à **l'hébergement web mutualisé**, ou aux **machines virtuelles (VM)**

- **Modération** des dépôts pour l'attribution du label données de qualité dans Nakala : accompagnement au dépôt et curation des données
- Recherche de ressources et mise en relation avec les **consortiums Huma-Num**

Plateforme Universitaire de Données en Bretagne (PUD-B)

- Accompagnement dans la conception et la réalisation de questionnaires d'enquête (en lien avec PACIHM)
- Accompagnement dans l'accès à des bases de données existantes (Quetelet-Progedo, enquêtes internationales, etc..)
- Accompagnement individuel de projets portant sur l'analyse statistique et économétrique de bases de données :
    - Utilisation de plusieurs logiciels
        - R / Rstudio
        - Python (Spyder, Visual Studio Code)
        - Jamovi, JASP
        - SAS
        - ARX (pour anonymisation)
    - Nettoyage de bases de données
        - Correction d'erreurs de saisie
        - Gestion des valeurs manquantes
- Suppression et/ou imputation de valeurs manquantes
    - Uniformisation de formats
        - Vérification de cohérence logique
        - Recodage des variables
- Regroupements d'informations pour gagner en puissance statistique
    - Exploration des différentes méthodes potentiellement utile au projet. But de pouvoir faire
        - Méthode simple (Statistique descriptive)
        - Méthode intermédiaire (analyses bivariées)
        - Méthode plus complexe (analyses multivariées)
    - Aide sur la manière dont chaque méthode s'analyse
- Possibilité de faire des demandes plus atypiques
    - Webscrapping
    - Préparation de jeux de données (traitement, anonymisation ou pseudonymisation, documentation (ajout de métadonnées)) en vue d'un dépôt dans un entrepôt (Quetelet-Progedo, Nakala, etc.)
    - Formation à la demande ou dans le cadre des DataLabs sur des logiciels (R, Jamovi, ARX, JASP) et thématiques statistiques et de gestion de données.

Plateforme d'Analyse du Comportement et des Interactions de l'Humain dans son Milieu (PACIHM)

Accompagnement à la collecte et l'utilisation de données sur les interactions humaines (quali, quanti) par la mise à disposition de ressources matérielles et l'expertise associée.

- Conseil sur les méthodes, outils et équipements de mesure et d'analyse (méthode expérimentale, observation, questionnaire, entretiens, eye tracking, EEG...) de données sur le comportement humain (physiologiques, comportementales, subjectives, auto-rapportées...).
    - Conseil méthodologiques (à l'élaboration du protocole, pour l'élaboration d'un questionnaire, échantillonnage, choix des mesures...)
    - Conseil sur les équipements (conseils à l'achat, mise en relation avec structure possédant le matériel envisagé...)

- Soutien technique en analyse du comportement humain :
    - Mise à disposition de ressources matérielles
        - équipements (matériel et logiciels) :
            - de simulation (e.g., simulateur de conduite, casque VR)
            - de mesures physiologiques (EEG, eye tracking...)
            - de mesures comportementales (eye tracking écran, eye tracking mobile [glasses], système d'enregistrement vidéo...)
        - locaux (salle d'expérimentation partagée, bureau),
        - base testeurs (panel > 500 participants pouvant être sollicités pour des études en ligne ou sur site)
    - Sensibilisation/formation sur les méthodes et à l'utilisation des équipements (mise à disposition de documentation et/ou formation à la demande)
        - Prise en main des équipements de la plateforme (voir liste ci-dessus)
        - Prise en main de logiciels :
            - The Observer (logiciel d'acquisition et de synchronisation de données)
            - Tobii pro Lab (logiciel d'analyse de données eye tracking)
            - STISIM (logiciel de simulation de conduite)
            - NVivo (logiciel d'analyse de données qualitatives) (à venir)
            - Psychopy (logiciel de conception d'expérimentations) (à venir)
            - Limesurvey (outil d'élaboration de questionnaires en ligne)
    - Sensibilisation/formation au traitement des données récoltées sur la plateforme (données eye-tracking, simulateur de conduite...) (mise à disposition de documentation et/ou formation à la demande) (selon équipements concernés)
    - Accompagnement à l'utilisation des équipements, au recueil des données, à la mise en place des expérimentations (e.g., appui passations, coordination de stagiaires...)