**Pratiques éthiques pour les annotateurs (version 2020-02-28)**

Base du document préparé par : Djegdjiga AMAZOUZ [djegdjiga.amazouz@gmail.com](mailto:djegdjiga.amazouz@gmail.com)

Edité par Alejandrina Cristia

1-Nom et coordonnées de la personne responsable du corpus :

Alejandrina Cristia (Chercheur principal) [alecristia@gmail.com](mailto:alecristia@gmail.com)

2-Description:

Ce document de pratiques éthiques pour les annotateurs présente des lignes de description de données à annoter, les éventuelles situations auxquelles l'annotateur peut faire face au moment du traitement des données et la mise à disposition des données pour l’usage académique. Les éléments prévus dans ce guide de l’éthique sont les différentes réglementations et conduites d’annotation à suivre par les annotateurs, la réglementation relative à la propriété intellectuelle des données et la contribution intellectuelles des annotateurs.

3-Résumé des tâches:

Vous allez annoter des données provenant d'enregistrements de personnes dans leur vie quotidienne avec au moins un jeune enfant. Ces clips ne peuvent pas être considérés comme "anonymes" car les voix n'ont pas été altérées et des métadonnées concernant les intervenants ont été recueillies. De plus, les enregistrements peuvent être accompagnées des fonctions corporelles occasionnelles, ou des moments intimes (désaccord ...).

Bien que nous possédions un consentement des principaux participants aux enregistrements, certains des participants secondaires inclus dans ces enregistrements n'ont pas donné leur accord d’enregistrement et d’utilisation des données. En revanche, comme les enregistrements sont des rencontres quotidiennes, ils peuvent contenir des paroles de personnes dont nous n'avons pas directement la permission d'utiliser la voix (par exemple, un étranger dans un parc, un caissier dans une épicerie). Soyez particulièrement soucieux dans ces moments ! S'il vous semble qu'ils ne sont pas au courant d'être enregistrés, ne continuez pas à écouter.

3-Réglementation et conduite d’annotation:

3.1. La copie des données et des données annotées sur un stockage personnel (ordinateurs, stockage externe, téléphone, drive en ligne ...) n’est pas autorisée. Seules les copies dans votre espace électronique du laboratoire sont admises (espace dédié à l’annotation).

3.2. Les personnes enregistrées doivent être respectées. Les annotateurs doivent maintenir donc la confidentialité des données. Les faits perçus et les conversations écoutés dans les enregistrements (parole, voix et les divers sons) ne doivent pas être communiqués aux personnes qui sont en dehors de l’équipe de recherche. Aussi, Il n’est pas autorisé de retranscrire, ré-enregistrer, photographier les données écoutées ou lues pour une rediffusion publique.

3.4. Soyez particulièrement attentif lorsque vous entendez des informations personnelles (noms, adresses, mails, informations sur les cartes de crédit ... ) car nous ne voulons pas analyser ces morceaux. Spécifiquement, signalez avec "pas écouter" des morceaux avec :

* des informations qui peuvent identifier une personne (nom complet, numéro de téléphone, adresse, numéro de sécurité sociale, date de naissance
* un épisode qui semble privé (sons sexuels)
* des commentaires sur tierces personnes qui peuvent être mal pris ("XX est un imbécile complet")
* des échanges qui semblent indiquer une activité illégale
* des disputes graves

3.5. Ne faites pas une écoute biaisée concernant l'aspect culturel de la parole et des actes (écoute sans sensibilité culturelle). Même si la famille qui est enregistrée semble appartenir à notre culture, on ne peut pas être sûrs de comprendre ses propres règles internes et ses modes de comportement (relations et comportement entre les membres de la famille).

3.6. Nous ne sommes pas tenus à révéler si nous avons entendu parler d'activités illicites dans les enregistrements, telles que la consommation de drogue ou la violence domestique. Si avoir entendu ceci vous a affecté, voir le point suivant.

3.7. Il est demandé aux annotateurs de garder une distance psychologique par rapport à ce qu'ils entendent et de ne pas porter de jugement d’ordre social ou culturel. Si l'annotateur est néanmoins troublé ou gêné par un passage qu'il entend, il doit en parler avec le chercheur principal et ne pas en discuter avec des tiers, car cela pourrait violer la vie privée et la confidentialité du participant. Le chercheur principal recommandera alors l'accès à un soutien psychologique externe et/ou à une discussion en groupe de pairs avec les autres membres de l'équipe.

4-Propriété intellectuelle:

4.1. Les corpus, en particulier transcrits et annotés, sont indispensables pour la recherche et les études. Les corpus mobilisent donc des ressources humaines et financières importantes.

4.2. La collecte de données, la génération des questions de recherche, l'élaboration des principes et matériaux d'annotation, et la coordination de l'annotation restent sous la responsabilité d'Alejandrina Cristia, qui décidera sur qui seront les auteurs du corpus en suivant les principes de l'APA (<https://www.apa.org/research/responsible/publication/>).

4.3. Votre travail dans notre équipe sera typiquement dans le cadre d'un stage aux fins éducatifs. Il est donc ne pas attendu que vous contribuez à la hauteur nécessaire pour devenir un co-auteur. Il est possible, néanmoins, que votre travail soit reconnu (dans une section "acknowledgments"), si vous en êtes d'accord.

4.4. Vous serez autorisé à utiliser ses données et des données dérivées dans des objectifs éducatifs pré-définis. Néanmoins, à la fin de votre stage, vous n'auriez plus accès aux données (voir section 3.1).

Annexe: Quelques définitions

Donnée : toute information constitue potentiellement une donnée. Une information devient une donnée dès lors qu’elle est recueillie.

Donnée brute : toute donnée recueillie et n’ayant subi aucun traitement algorithmique ou manuel.

Donnée annotée: toute donnée à laquelle des annotations manuelles ou automatiques ont été attribuées.

Annotateur: tout individu qui intègre l’organisation et chargé d’écouter et d’annoter les données, peu importe le statut sous lequel il est désigné contractuellement (CDI, CDD, stagiaire, apprenti, intérimaire, bénévole, etc...).

Corpus: ensemble homogène, significatif et organisé de données (sons, images, écrit …) observées et à partir desquelles pourra effectuer des recherches et s'élaborer la théorie.

Bonne pratique : tout ensemble de comportements qui font consensus et qui sont considérés comme indispensables par la plupart des professionnels pour la qualité d’un domaine donné. La gestion des bonnes pratiques dans le domaine des sciences humaines et sociales implique une veille stratégique et une gestion approfondie des connaissances de ce domaine.