

Marie-Aude Fouéré (IFRA Nairobi) et Éloi Ficquet (EHESS-CEIFR)

**Greniers numériques et termitières du web :
méthodes de collecte et de conservation des données d'internet en contextes
africains**

Les études en sciences humaines et sociales sur les TIC s'intéressent le plus souvent à leurs usages actuels et la manière dont elles transforment les pratiques communicationnelles et les sociabilités, mais également influencent les phénomènes politiques d'aujourd'hui et contribuent à l'émergence de subjectivités nouvelles. Cependant, les réflexions d'ordre méthodologique sur la collecte et le traitement des données issues des TIC sont singulièrement absentes de ces travaux. Comme pour tout autre terrain d'enquête, il est indispensable d'explicitier les procédures mises au point pour échantillonner et de constituer un corpus de données. Les contenus foisonnants que le chercheur doit découper, sélectionner et ordonner doivent être analysés en lien avec ses producteurs ou émetteurs et ses récepteurs, ainsi qu'avec les modalités et les réseaux de leur mise en circulation.

Cette communication porte sur les méthodologies employées pour la collecte et de l'analyse de champs de données d'Internet (blogs, YouTube par exemple) et aux autres dispositifs d'auto-communication de masse (Facebook, Twitter). Il s'agira de voir dans quelle mesure des techniques des sciences humaines déjà éprouvées sur d'autres terrains et objets peuvent être adaptées à des corpus en expansion et de plus en plus sollicités comme données complémentaires ou principales d'une enquête. Dans quelles conditions faire de l'observation ethnographique sur ces sociabilités nouvelles ? Quelle critique philologique peut être appliquée à des documents constitués de strates d'information hétérogènes et volatiles ?

Ces questionnements généraux peuvent se poser avec des inflexions particulières quand ils portent sur des corpus numériques produits ou circulant principalement dans des sociétés africaines. Par exemple, l'usage des langues africaines, souvent mêlées à des langues de communication internationale, interrogent les procédures de traitement semi-automatique de l'information. Les obstacles techniques et procéduraux à l'hébergement de sites et de serveurs donnent aux communautés émigrées des rôles décisifs, voire excessifs, de relais d'information. On pourra aussi interroger les stratégies d'évitement, de contournement de dispositifs de censure.