

## Proposition de sujet de Thèse

### a. Titre de la Thèse :

Analyse des données issues des réseaux sociaux africains afin de déterminer des pistes de développement de l'Afrique à travers les besoins formulés sur ces réseaux.

### b. Directeur(s) de Thèse(s) :

Awa Diattara	awa.diattara@ugb.edu.sn	Université Gaston Berger, Saint-Louis, Sénégal
Assitan Traoré	assitan.traore@actoll.com	ACTOLL, Grenoble, France

### c. Description du sujet de thèse

#### i. Contexte

Les réseaux sociaux représentent une mine d'informations qui peuvent être exploitées pour identifier des projets d'amélioration de la vie quotidienne des africains [5]. Avec l'apparition et la vulgarisation des réseaux sociaux depuis 2006 dans le monde, ils sont aujourd'hui devenus incontournables dans la vie quotidienne pour des millions de personnes pour échanger, établir des liens mais aussi pour parler des préoccupations ou partager son indignation face aux événements du quotidien. Ces réseaux représentent donc un forum géant international de discussion qui mérite d'être exploité pour l'identification et la mise en place de projets innovants de développement. Plusieurs travaux de recherche existent aujourd'hui exploitant ces données issues des réseaux sociaux pour la découverte de nouvelles connaissances dans des domaines divers comme le terrorisme (analyse des données des réseaux sociaux pour la détection de la radicalisation des gens [6]), la cybercriminalité (la prédiction du genre et l'âge en fonction des interventions sur les réseaux sociaux [7]), etc.

#### ii. Objectif

L'idée principale de cette thèse est de faire une analyse de plusieurs réseaux sociaux africains avec les méthodes d'IA et de fouille de donnée (text mining, traitement d'image) pour définir des besoins qui seront par la suite formulés en projet de développement pour répondre à ces besoins [1] [2] [3]. La problématique de Big Data sera abordée dans cette thèse pour l'accès, la collecte et exploitation de ces données issues de plusieurs réseaux sociaux [4].

#### iii. Méthodologie

Cette thèse sera constituée de trois grandes étapes qui sont :

1. L'étape de collecte des données issues des réseaux sociaux : selon la disponibilité des données des réseaux sociaux de préférence africains. La collecte se fera au minimum sur deux réseaux sociaux. Le doctorant devra travailler pour identifier ces réseaux et trouver un moyen d'accès aux données de ces réseaux.
2. L'étape d'analyse des données collectées : les données collectées devront être prétraitées (préparées pour l'analyse), analysées avec les méthodes de text mining et de traitement d'images pour identifier des mots clés indicateurs de besoins des usagers de ces réseaux.
3. La formulation des projets de développement à destination de l'Afrique : une fois identifiés, les mots clés seront étudiés pour définir des projets de développement durable construits à partir des besoins formulés sur les réseaux sociaux.

#### **iv. Références bibliographiques**

1. Jana Diesner, Chieh-Li Chin, Marc A. Smith: Combining Online Social Networks with Text Analysis. Encyclopedia of Social Network Analysis and Mining. 2nd Ed. 2018.
2. Text Mining. Encyclopedia of Social Network Analysis and Mining. 2nd Ed. 2018.
3. Text Networks. Encyclopedia of Social Network Analysis and Mining. 2nd Ed. 2018.
4. Sancheng Peng, Shui Yu, Peter Mueller: Social networking big data: Opportunities, solutions, and challenges. Future Generation Comp. Syst. 86: 1456-1458 (2018).
5. Sancheng Peng, Guojun Wang, Dongqing Xie: Social Influence Analysis in Social Networking Big Data: Opportunities and Challenges. IEEE Network 31(1): 11-17 (2017).
6. S. Agarwal et A. Sureka, « Applying Social Media Intelligence for Predicting and Identifying On-line Radicalization and Civil Unrest Oriented Threats », ArXiv151106858 Cs, nov. 2015.
7. Peersman, W. Daelemans, et L. Van Vaerenbergh, « Predicting age and gender in online social networks », in Proceedings of the 3rd international workshop on Search and mining user-generated contents - SMUC '11, Glasgow, Scotland, UK, 2011, p. 37.

#### **d. Profil recherché du doctorant**

- Master en Informatique (ou diplôme équivalent)
- Des compétences dans l'un des domaines : IoT, Big Data, Machine Learning ou analyse de données seront appréciées.



**UNIVERSITÉ GASTON BERGER SÉNÉGAL**

CENTRE D'EXCELLENCE AFRICAIN SUR LES MATHÉMATIQUES,  
L'INFORMATIQUE, ET TIC (MITIC)



**PASET**

Partnership for Skills  
in Applied Sciences,  
Engineering & Technology

FONDS RÉGIONAL DE BOURSES D'ÉTUDES  
ET D'INNOVATIONS (RSIF) DU PASET

## PhD Research Proposal

### a) Subject's Title :

Analyzing african social networks data to identify development tracks for Africa through concerns formulated on african social networks.

### b) PhD Supervisor(s):

Awa Diattara	awa.diattara@ugb.edu.sn	Université Gaston Berger, Saint-Louis, Sénégal
Florent Devin	florent.devin@eisti.eu	CY TECH (ex EISTI), Pau, France
Assitan Traoré	assitan.traore@actoll.com	ACTOLL, Grenoble, France

### c) PhD Description

#### i) Context

Social networks represent a wealth of information that can be used to identify projects in order to improve the daily lives of africans [5]. Since their appearance and popularization around the world in 2006, social networks have become essential in daily life of millions of people to exchange, establish link but also to talk about concerns or share their indignation against everyday events.

These networks therefore represent an international giant forum for discussion that deserves to be exploited in order to the identify and implement innovative development projects. Today several research works exist and use data from social networks to discover new knowledge in various fields such as terrorism (analysis of data from social networks to detect the radicalization of people [6]), cybercrime (prediction of gender and age based on social media interventions [7], etc).

## **ii) Research Objectives**

The main idea of this thesis is to make an analysis of several African social networks with AI and Data Mining methods (text mining, image processing, etc.) to identify needs that will be formulated as development projects to meet these needs [1], [2], [3]. The Big Data issue will be addressed in this thesis for the access, the collection and the analysis of social networks data.

## **iii) Methodology**

This thesis will consist of three main steps which are:

1. The collection of data from social networks: depending on availability of data on social networks, preferably African social networks. Data collection will take place on at least two African social networks. The PhD student will have to investigate to identify these networks and find a way to access data.
2. The analysis of data collected: the collected data must be preprocessed (prepared for analysis) and must be analyzed with text mining and image processing to identify keywords that indicate the concerns of social network users.
3. The proposition of development projects for Africa: once identified, the keywords will be studied to define sustainable development projects built from social networks user's needs.

## **iv) References**

1. Jana Diesner, Chieh-Li Chin, Marc A. Smith: Combining Online Social Networks with Text Analysis. Encyclopedia of Social Network Analysis and Mining. 2nd Ed. 2018.
2. Text Mining. Encyclopedia of Social Network Analysis and Mining. 2nd Ed. 2018.
3. Text Networks. Encyclopedia of Social Network Analysis and Mining. 2nd Ed. 2018.
4. Sancheng Peng, Shui Yu, Peter Mueller: Social networking big data: Opportunities, solutions, and challenges. Future Generation Comp. Syst. 86: 1456-1458 (2018).
5. Sancheng Peng, Guojun Wang, Dongqing Xie: Social Influence Analysis in Social Networking Big Data: Opportunities and Challenges. IEEE Network 31(1): 11-17 (2017).
6. S. Agarwal et A. Sureka, « Applying Social Media Intelligence for Predicting and Identifying On-line Radicalization and Civil Unrest Oriented Threats », ArXiv151106858 Cs, nov. 2015.
7. Peersman, W. Daelemans, et L. Van Vaerenbergh, « Predicting age and gender in online social networks », in Proceedings of the 3rd international workshop on Search and mining user-generated contents - SMUC '11, Glasgow, Scotland, UK, 2011, p. 37.

#### **d) Profil Candidate proproposal**

- Research Master Degree in Computer Science (or equivalent diploma)
- Skills in one of the fields: IoT, Big Data, Machine Learning or data analysis will be appreciated