



Semantics for Trust – Sem4Trust

Approche sémantique pour l'étude
des échanges sur les réseaux sociaux

3 équipes du Département IA de l'IRIT

Recrutement d'un IR durant 1 an : Pierre-Yves Gicquel

Pascale Zaraté (PR UT Capitole, ADRIA),

Yannick Chevallier (MCF UT, LiLAC) et Nathalie Aussenac-Gilles (DR CNRS, MELODI)

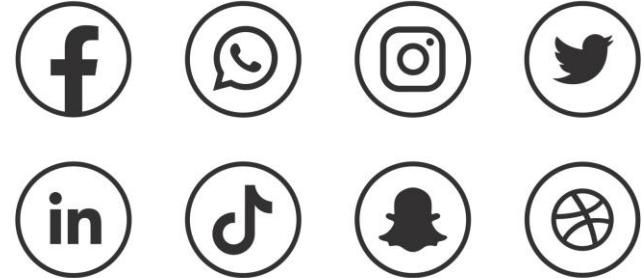
Prenom.nom@irit.fr



L'étude des réseaux sociaux : enjeux pour la cyber sécurité



- Différentes natures de cybercriminalité
 - Pédophilie, traite de personnes
 - Diffusion de fausses informations, de rumeurs
 - Diffamation, harcèlement, atteintes aux personnes
 - Incitation à la haine, racisme
 - Préparation d'actions violentes, d'attaques terroristes
 - Influence en vue de recrutement de personnes (sectes, groupes terroristes ..)
- Différentes formes de cybercriminalité
 - Diffusion de contenus illégaux
 - Diffusion en masse d'idées et de messages de désinformation
 - Usurcation d'identité, utilisation d'identités multiples
 - Participation à des conversations dans des buts malveillants
- Thématique ICO : **Mieux protéger les données et la vie privée, et améliorer la confiance dans les réseaux sociaux.**



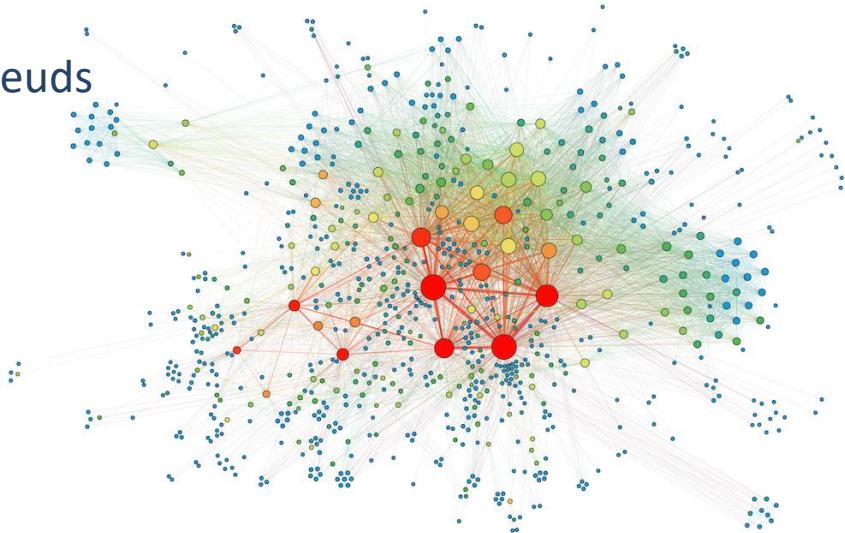
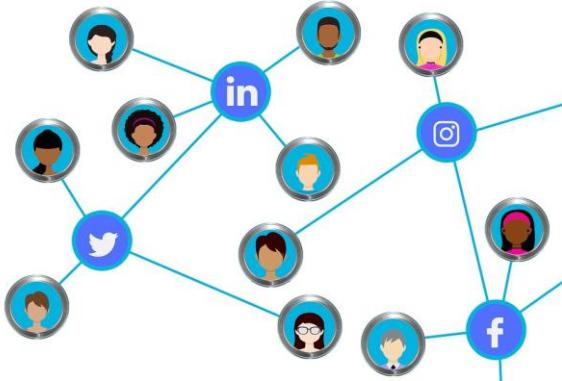
L'étude des échanges sur les réseaux sociaux : une approche réseau

- **Graphe des échanges**

- Chaque site et chaque contributeur est un nœud.
- Les relations entre sites, entre contributeurs, etc correspondent aux arcs.

- **Nature des études**

- Importance et nature des flux
- Volume des consultations (popularité) des noeuds
- Forme et structure du graphe



<http://www.martingrandjean.ch/introduction-visualisation-de-donnees-analyse-de-reseau-histoire/>

L'étude des échanges sur les réseaux sociaux : étude des contenus

- Images, vidéos, son, langage naturel écrit ou oral

- Traitement Automatique de la langue naturelle écrite
 - Contributions individuelles
 - Conversations et des échanges entre contributeurs
 - Étude des thématiques, des tendances, des influences
 - Recherche de comportements suspects, sexistes, de la production et de la propagation de rumeurs et de fausses informations

L'étude des échanges sur les réseaux sociaux : étude des éléments de contexte

- Contexte de production des contributions
 - Localisation, volume, langue, date
 - Caractérisation et évolution dans l'espace et le temps des contributions
 - Contributeurs
- Où trouver ces éléments ?
 - Méta-données des sites, des messages, des posts
 - Traces de consultation, de navigation, cookies
- Qui a accès à ces éléments ?
 - Les entreprises / organismes qui gèrent ces réseaux
 - Les publicitaires, ceux qui achètent ces données
 - Difficiles pour les enquêteurs ou les chercheurs (RGPD, respect de la vie privée ...)

Expérience de l'IRIT dans

~~STARLIGHT~~

Sustainable Autonomy and Resilience for
LEAs using AI against High Priority Threats

- **STARLIGHT**

- Projet H2020 de 2022 à janvier 2026
- 3 équipes du Dept IA de l'IRIT + juristes UTC
- Solutions à base d'IA « éthiques et respectueuses de la loi » par conception

- **Focalisation IRIT**

- **Identification des comptes manipulés par un même individu**
- Exploiter métadonnées et contenus textuels
- Représentation sous forme de graphe
- Apprentissage mais ni réseaux neuronaux ni LLM
- Caractéristiques à tester pour chaque réseau social
- Validation à l'aide d'alter-égos



L'étude des échanges sur les réseaux sociaux : La plateforme logicielle de **Semantics4Trust**

- **Principes**

- Évolution du prototype réalisé pour Starlight
- Modularité : analyse des contributions, visualisation des résultats
- Représentation interne par fichiers CSV
- Fiable : validation des critères de comparaison par les utilisateurs
- Évolutif : pouvoir ajouter de nouveaux critères d'analyse

- **Données analysées**

- Méta-données des messages et des contributeurs (alias) : date, lieu
- Contributions en langage naturel
- Graphe des contributions

- **Algorithmes d'apprentissage testés**

- MLP, Random Forest et Regression Logistique

L'étude des échanges sur les réseaux sociaux : La plateforme logiciel de **Semantics4Trust**

- **Caractéristiques des contributions implémentées**
 - Volume des contributions par heure
 - Mots-clés les plus fréquents - mots rares les plus fréquents
 - Séquences de 1, 2 ou 3 lettres - Séquences de 1, 2 ou 3 mots
- **Fonctionnalités**
 - Importer un ensemble de conversations dans un graphe (CSV)
 - Découper le corpus en alter-égos
 - Entraîner le module d'analyse selon une des caractéristiques
 - Évaluer le module
 - Comparer des contributeurs selon cette caractéristique
 - Choisir et combiner plusieurs caractéristiques de comparaison
 - Regrouper les contributeurs similaires selon les caractéristiques choisies

L'étude des échanges sur les réseaux sociaux : La plateforme logiciel de **Semantics4Trust**

- Ingénieurs recrutés

- . Pierre-Yves Gicquel, 10 ans d'expérience en post-doctorat et IR Spécialiste en représentation des connaissances et web sémantique ; recruté au CNRS comme IR au bout de 8 mois
 - . Khalil Maachoud, 2 ans d'expérience, M2 ingénieur IA de l'UT

- Plateforme

- Prochainement disponible sur Gitlab IRIT
 - Test en cours avec la police Tchèque



Merci de votre attention

