

# GEOL Knowledge

GEOL Knowledge semantic profiling identifies the actors cited in texts and their personal data

Thanks to the SAIMSI project, funded by ANR, GEOL Knowledge answers the questions: WHO? WHAT? WHEN? WHERE? HOW? HOW MANY? MUCH?

GEOL Knowledge transforms texts written in several languages (French, English, Arabic and Chinese) in language independent structured information according to an ontology describing the application.

To reach this goal, GEOL Knowledge uses a general purpose deep morphosyntactic analysis, a rule-based application-oriented semantic extraction and a final tool to infer additional knowledge from a global view of the document and its metadata.

An ontology and extraction rules are already available for personal information for security

and GDPR: attributes of person, family relations, contact, emission of a message, relation with police and justice, death, violent act, meeting, travel, job, education. In addition to SAIMSI results, GEOLSemantics has added for GDPR compliance : personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade union, health, sexual orientation... Insults are also extracted.

CONTENU DE LA BASE DE DONNÉES DES EMPLOYÉS		
EMPLOYÉ	DATE	COMMENTAIRES
Jean Dupont	3 février 2014	10 ans d'ancienneté
	15 septembre 2014	diabétique
	10 janvier 2015	feuille de paie transmise
	20 mars 2015	Rémard
	1 septembre 2015	contacter son fils au 09 10 11 12 13
	15 avril 2016	délogue syndical
	10 mai 2016	condamné pour excès de vitesse

  

RÉSULTAT DES ANALYSES	
DONNÉES IDENTIFIÉES	CLASSIFICATION
« 10 ans d'ancienneté »	données à caractère personnel conforme au traitement « gestion des employés »
« diabétique »	données sensibles – données de santé
« feuille de paie transmise »	pas de données à caractère personnel
« Rémard »	personnes impliquées
« fils » : « 09 10 11 12 13 »	données à caractère personnel
« délogue syndical »	données sensibles sauf exceptions définies par la réglementation
« condamné pour excès de vitesse »	données réservées aux autorités policières

## Extraction of personal data

**Key selling point(s):** Multilingual knowledge extraction, Language independent output in RDF, personal data, application for privacy

**Key market(s):** Security, Digital Cities, privacy protection, privacy by design, competitive intelligence, GDPR

**Main customer(s):** Ministry of interior and defence, large companies like Banks and insurances, all companies concerned by GDPR

**Systematic R&D project(s):** SAIMSI



## GEOLSEMANTICS

12, avenue Raspail  
94250 GENTILLY

[www.geolsemantics.com](http://www.geolsemantics.com)

**Contact:** Mr Christian Fluhr  
christian.fluhr@geolsemantics.com  
**Head count:** 8 employees