

Semantic Crime Investigation System

Eman K. Elsayed*, Wafaa Ghonaim, Eman Zeaar
Faculty of Science (Girl branch), Al-Azhar University, Cairo, Egypt.

* Corresponding author. Email: emankaran10@azhar.edu.eg
Manuscript submitted May 9, 2018; accepted July 10, 2018.
doi: 10.17706/jcp.13.10.1216-1226

Abstract: Fusion the direct and indirect evidences with crime knowledge base is important challenge in Automate the crime-solving processes. So, this paper proposed SCIS (Semantic Crime Investigation System), which supports investigator to take a certainty decision based on result of the fusion. The solution proposed in this paper (SCIS) used Ontology re-engineering to create Crime Universal Ontology which includes kinesics Ontology. Also SCIS applied sentiment analysis and image processing techniques. These concepts work together in SCIS in three phases to support the investigator by detailed collected report. The investigator can use SCIS as a mobile application. The proposed system is Automatic, mobile, general, scalable and fast system. Finally, in this paper there is a real case study.

Key words: Crime investigation, ontology based, image processing, sentiment analysis, kinesics.

1. Introduction

Criminology is considered as one of the most important fields that focus on studying of crime, criminal behavior and law enforcement. It aims to identify the characteristics of the crime. Crime-solving processes includes analyzing and discovering crime and determine the relations [1].

Criminal investigation has very important and significant role in any country and is a legal activity of state judicial organs to accomplish the task of criminal Law, detecting and collecting the criminal evidence, and investigating and arresting the criminals [2]. There exist many types of crime and criminal for example (cyber, killing, theft, Kidnapping, Rape and so on). There are many challenges for automation the crime solving processes and crime investigation as fusion the direct and indirect evidences to support the investigator decision. Also, the speed is more important so the proposed system in this paper (SCIS) Semantic Crime Investigation System is mobile application without the need of any specific devices or support from other people. We built SCIS based upon Image processing techniques, Body language analysis, text mining, pattern recognition and Ontology based. The characteristics of the complex and dynamic crime investigation domain require new ways of information extraction and knowledge representation [3]. So we used Ontology based in this paper, where Ontology rich with semantics than database [4]. One of the main contributions in ontology mapping is to grow the ability of concept linkages in order to gain a comprehensive knowledge among heterogeneous domains [5].

Knowledge base is the heavy Ontology i.e. Ontology with instances. It widely used in several domains to support intelligent systems. The development of application ontology helps to create the framework and thus to ensure the collection, accumulation, storage, treatment, and transmission, in a proper form, of important investigation information. Using Ontology has many advantages as sharing and common understanding knowledge in certain domain [6]. Also, Ontology improves recall by query expansion through the synonymy relations.

The aim of the article is creating crime investigation system based on universal knowledge based to help the investigator to discover the truth fast. In order to increase the efficiency and effectiveness of our system we merged the kinesics ontology in SCIS.

The remainder of this article is organized as follows: Section two introduces a brief survey on related researches. Ontology construction presents in section three. Section four presents the phases of the proposed SCIS (semantic crime investigation system) in details. Section five concludes our work and future works.

2. Review of the Related Research

There are many projects currently developed in crime investigation. All of them support different types of users to solve the crime or protect the societies from repeating the similar one.

In reference [7] authors proposed a crime investigation system against illegal smuggling in Taiwan by applying Case-Based Reasoning (CBR) and experts systems in conjunction with computer-assisted instruction. In their system investigators can input smuggling crime characteristics and approaches. CBR of smuggling crime characteristics and approaches is established which searches for most similar cases through case comparison for crime investigation. The similarity and weight of an attribute can be adjusted to enable the case inference mechanism of this CBR system to approximate the actual characteristics and approaches of the crime. Finally, concluded that the system provides for investigators comprehensive recommendations and helps them establish an overall investigation concept. But that helps only in certain type of crime called smuggling crime.

In reference [8], authors scope on crime analysis from the crime matching process. They satisfy that by using machine learning techniques neural network and K-means in two steps respectively. The training set is 21 crime variables. So we can classify this type of research as data mining research for extracting important entities from police narrative reports.

Also in reference [9], authors used Ontology based to propose a recommendation filling system. The crime scene investigation ontology was developed from information gathered by domain experts and assigned to Ontology expert in concepts, relationships, and definitions. The cases of 50 documents from the crime scene investigation report about properties of Forensic Science Police Center were analyzed.

Some references support crime-solving processes from pattern recognition perspective. That is by concerned about biometrics in crime scene. A biometrics-based crime investigation system for the Nigerian Police Force is proposed in reference [10]. That was to retrieve all existing criminal details based on the database and another processing task involves comparing Fingerprints. Suppose a criminals' fingerprint was recovered from a crime scene, the user loads this print into the system and searches for a match. Finally, a successful implementation of the Biometrics based crime records investigation system will greatly increase the efficiency of the Nigerian Police Force officers and will assist to confirm that criminal information records are managed properly monitoring of criminal suspects in the country.

Some researchers focus on the relation between the members of a criminal group by using a graph theory. Reference [11] supports crime investigation system by using graph theory, which clarifies the relative importance of nodes in a graph with respect to a given set of query nodes. In such a graph, a member of a criminal group represents as a node, the relationship between two members of the group represents with an edge connecting two nodes, and the degree of the relationship between those two members represents by the weight of an edge. By using this depiction, the authors determined the relative importance of nodes in a graph with respect to a given set of query nodes. Most present methodologies that study relative significance determine it for a node under consideration by evaluating the contribution of each query node separately to the importance of this node while overlooking the contribution of the query nodes cooperatively to the importance of the node under thought. Three different datasets are used for the evaluation.

Ref. [12] also proposed SGEDSS (semantic gene expression decision support system). SGEDSS supports users to discover knowledge from heterogeneous data in communication networks (telephone numbers and customer databases). The application in [12] helped to follow up the suspects.

The contribution of this paper in crime-solving processes is general overview by merging direct and indirect evidences via Ontology re-engineering.

3. Construction of the Universal Ontology for SCIS

In this article, we support the proposed system SCIS by Universal Ontology Fig. 1. It consists of two main parts proposed ontology (POCIS) and kinesics ontology (KO).

We used Protégé [13] as OWL Ontology building platform. Querying the ontology's classes and instances is an important and essential service once ontology has been created. Querying process involves more semantic and logic operations. The queries are expressed by SPARQL on OWL (Ontology web language).

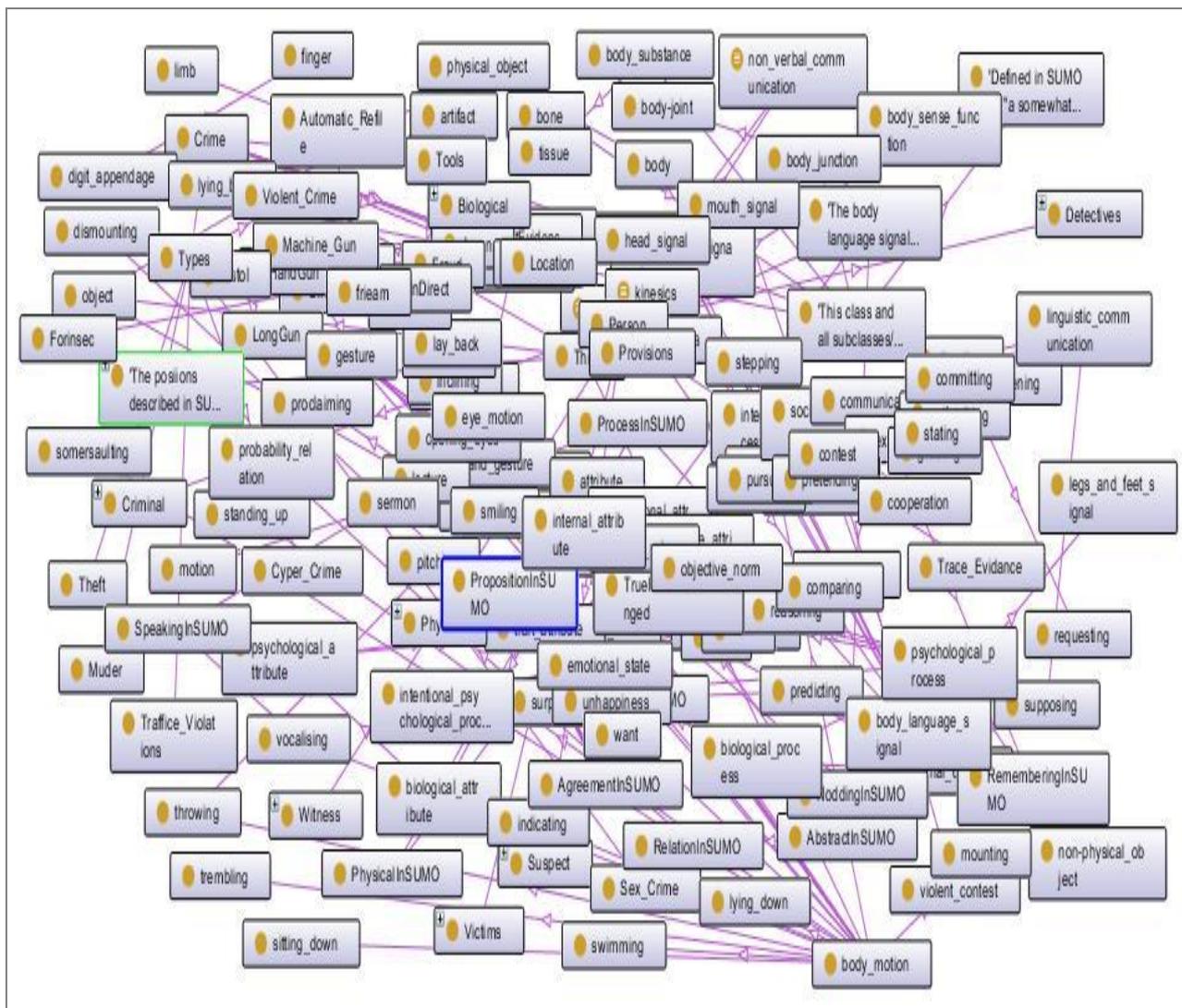


Fig. 1. The universal ontology visualization, it consists of two main parts proposed ontology (POCIS) and kinesics ontology (KO).

First part the proposed ontology (POCIS) which designed based on Egypt Civil Registry database information. In POCIS we focus on the crime concepts, relationships and axioms which are likely to be relevant to the criminal investigation domain. Based on a top-down approach we identify root class first,

subclass, the properties of classes, and object property between individual. The structure of the POCIS consists of five main classes (*Crime, Execution Area, person, Tools, Low, and Provisions*). For the *Crime* class which is the main class for all the crime concepts that are defined as three subclasses: Aim, Evidence, and Types. Every crime individual has properties Crime_Id, Crime_date, Method, Crime_tool, Crime_mage, Crime_Locaton, Crime_type and Finger_prints that exist in crime scene. For *person* class which is the main class of all crime person concepts, this class has five subclasses: Criminal, Detectives, Suspect, Victims, and Witness. Every person individual has Id, name, fingerprint, age, address, sex, religion and nickname. Criminal image and criminal type are additional individuals for Criminal subclasses. For *Tools* class which is the main class of all tools concepts, this class has two subclasses: direct and indirect. For the individual of direct subclass has Tool_Id as data property. For the four properties were used to assert axioms: has_tool relates <crime> [domain] with <Tools> [range], has_crime relates <person> [domain] with <Crime> [range], has_weapon relates <person> [domain] with <Tools> [range], islocatedIn relates <tools> [domain] with <crime> [range]. This general design can connected directly with any database. Also it can relate the crimes and criminal with each other.

Second part is Kinesics ontology (KO). Kinesics (body language) is a type of non-verbal communication and also is a non-verbal behavior related to movement of the body includes eye movements, facial expressions, posture, gestures, and the like. Lying is often more cognitively demanding than truth telling so, for the police interviews with real-life suspects lies were accompanied by decreased blinking, increased pauses, and decreased hand and finger movements, all of which are signs of cognitive load. Also suspects appeared to be thinking harder when they lied than when they saying the truth. For example, eyes tend to look right when the brain is imagining or fabricating, guessing and lying. But it looks left when the brain is remembering or recalling or retrieving something. The example in the case study in this article focus on the eye signals, eyes are a very important part of the non-verbal signals we send to others. Kinesics ontology contains 15 classes and 319 instances [14]-[16]. Sample of it was illustrated in Fig. (2a), Fig. (2b).

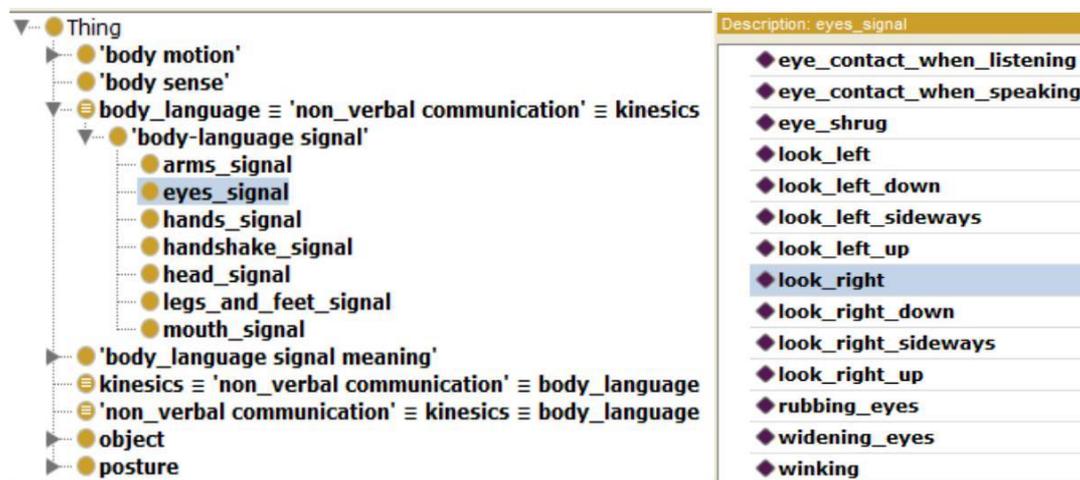


Fig. 2. (a) The class hierarchy of kinesics ontology representation and sample of instances of the eyes signal class is presented.

4. The Proposed Semantic Crime Investigation System (SCIS)

The proposed system architecture is illustrated by using UML sequence diagram as shown in Fig. 3. The UML sequence diagram presents the phases and their steps in details. The proposed SCIS has been implemented using MATLAB 2015 platform, SPARQL for queries and Protégé for OWL Ontology. SCIS consists of three phases: Crime scene, investigation, and final report phase. The details present in the following

subsections.

4.1. Crime Scene Phase

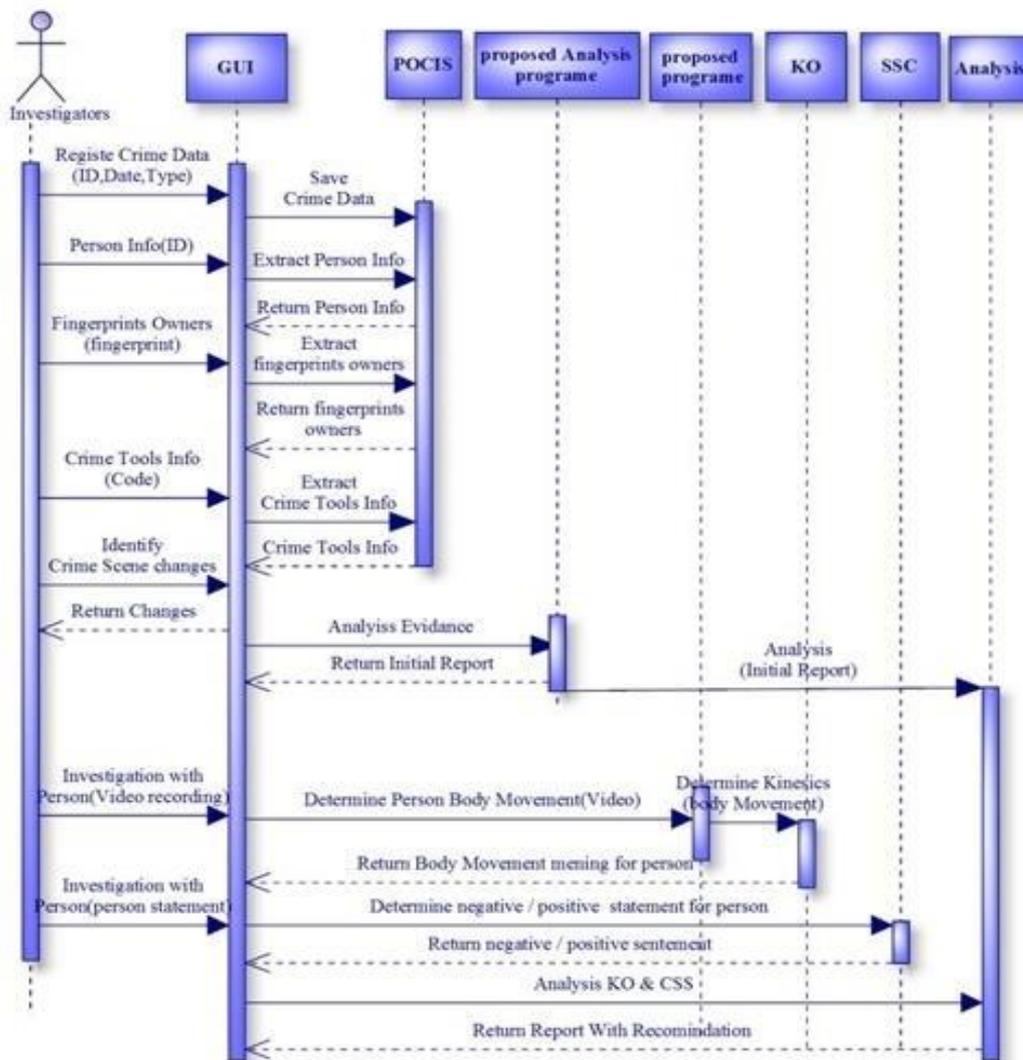


Fig. 3. UML sequence diagram of SCIS.

This phase is a main phase for identifying, collecting, and analyzing direct evidence. It consists of three different steps.

1. Crime registration step: This is the first step in which an important information about the crime is register may be crime number, date, crime type, investigator name, crime location , and crime tool if exist. The crime information inserts in the Universal Ontology.
2. Direct evidences collection step: the forensic evidence as (Fingerprints, blood traces, spent shells, saliva, semen, hair, fiber) and some pictures and Camera recordings in crime scene are collected manually, then insert also in the Universal Ontology. Also all persons (accused/suspects) information and crime tool information may be weapon number also collected. There are other related evidences in the crime scene called mock evidence as suicide note, medication pen, and glass.
3. Evidences analysis step: In this step we analyze the collected evidence automatically based on relations in POCIS Ontology part and the proposed program as the following:
 - Analyzing the persons (accused/suspects) information, that is by using the SPARQL queries as shown in Fig. 4 and using POCIS classes (person class). The search results (main person

information) are displayed and the subclasses name (Criminal or Detectives or Suspect or Witness) are retrieved to know which the subclass the person belongs to. The object property 'has_crime' is used to know the 'crime_Id' also by using 'has_weapon' as the object property we can retrieve if this person has weapon or not, if yes retrieve ' weapon_Id'.

Also vise verse, for crime tool information the user enter the weapon number for the proposed system, based on the POCIS class (Tool class) and object property (has weapon) the main weapon information are retrieved and the weapon owner is defined.

- Image processing: In order to automatically identify changes in crime scene objects we use the change detection proposed algorithm to extract all the tiny changes in the crime scene, the images are automatically recognized to identify objects. As illustrated in Fig. 5 we can compare the crime scene before and after crime to detect the changes in crime scene, may be changing the position of object or new objects are inserted.

Example in our case study, we could detect the existence of new object (as insect) in the crime scene. This maybe means that someone opens the window.

- Fingerprint matching: After collecting crime scene fingerprints, the user can compare features of fingerprints that are collected from the crime scene with the fingerprints features that are stored as data property values for individuals of person class. Based on SPARQL query that is shown in figure (6) and using data property (fingerprint) the fingerprint owner is retrieved. Also vise verse we can retrieve the name of subclass which the person belong to (e.g. criminal or Detectives or Suspect or Witness). Also we can retrieve all information (Id, name, nickname, age, address, religion, sex, type and image if exist) for this person using data property, we can retrieve crime_Id if this person sharing in the crime befor. Also we can retrieve Weapon which is owned by this person if exist. All the analysis outputs are saved automatically.

```
Select? Text
Where {
?sub ontoCrime2: Id? obj.
FILTER (regex (str (? obj)," id "))
?sub ontoCrime2: Data? text}
```

Fig. 4. Person information SPARQL query.

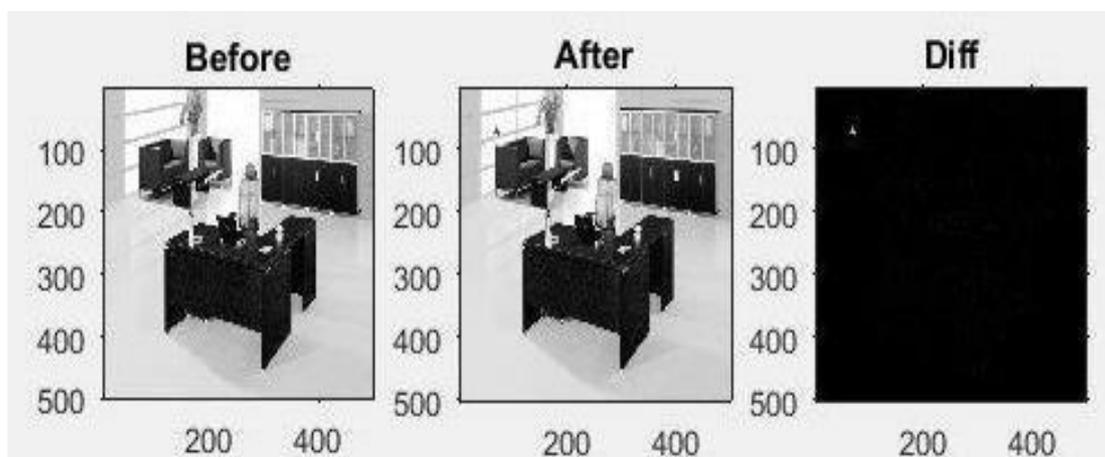


Fig. 5. Change detection in crime scene.

```
Select ?Individual
Where {
?Individual ontoCrime2:fingerprint ?fingerprint.
FILTER (regex(str(?fingerprint), "'fingerprint_features'")) }
```

Fig. 6. Fingerprint information SPARQL query.

4.2. Investigation Phase

According to the primary information that extracts from analysis crime scene phase the investigation phase is performed. During the investigation indirect evidences are collected based on both Kinesics Ontology and statement sentiment classification to know the attitude of a speaker (witness, suspect and accused) during the investigation. This phase is illustrated as the following:-

- *Kinesics ontology* (KO): in this phase SCIS uses KO to analysis body language. In this article we apply the system on the eye signal as indicator of speaker behavior by using look_left and look_right as a sample of eye signal instance. Using MATLAB code with mobile camera we can automatically detect eye signal (look_left or look_right) by comparing the distance from left edge to center point with the distance from right edge to center point. The result of our sample is displayed in Fig. 7. Based on Kinesics Ontology the object property of 'look_right' and 'look_left' individuals are shown in Fig. 8. The KO properties were used to extract axioms as in the following example: 'hasPossibleMeaning' relates in <body language signal> [domain] that have (e.g. look_left) as the individual for subclass of [domain] with <body language signal meaning> [range] that have (e.g. remembering, recalling, retrieving facts) as the individuals for subclass of this [range]. Also, 'hasPossibleMeaning' relates in <body language signal> [domain] that have (e.g. look_right) as the individual for subclass of [domain] with <body language signal meaning> [range] that have (e.g. lie, guessing, storytelling, Creating, fabricating) as the individuals for subclass of this [range]. SCIS automatically determined eye signal, then used a SPARQL query as shown in Fig. 9 to retrieve all possible meaning.
- Statement Sentiment Classification (SSC): The sentiment analysis is a complicated problem which divided into three subtasks: sentiment classification, subjective text detection, and subjective information extraction [17]. Sentiment analysis features show exactly how speakers (witness, suspect and accused) feel about discussion, and telling the investigator whether a given document of text is positive, negative, or neutral. During the investigation phase the statements of speakers are analyzed and classified into positive or negative sentiment, or neutral. The classification is performed based on sentiment analysis using Toolkit (NLTK) powered text classification process [18]. NLTK is free, open source, and easy to use interface which classify the text of different language. As illustrated in Fig. 10, Fig. 11, and Fig. 12 the sentiment analysis results of a set of statements.

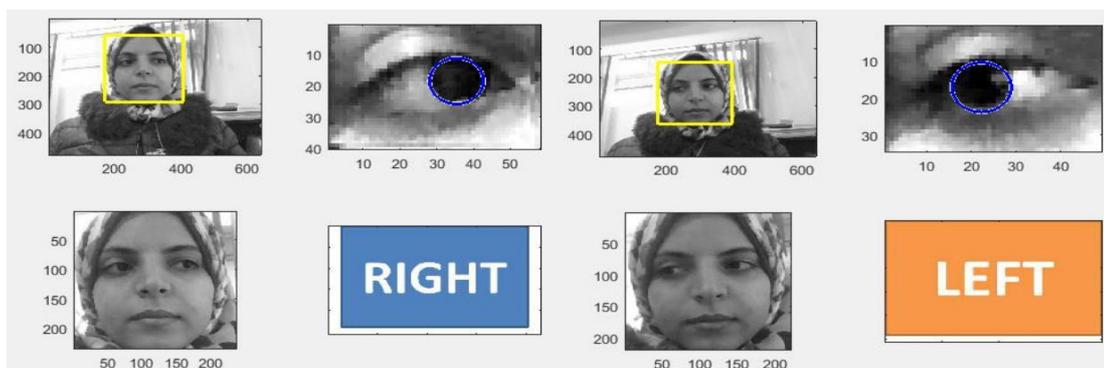


Fig. 7. Detect eye signal (look_right or look_left).

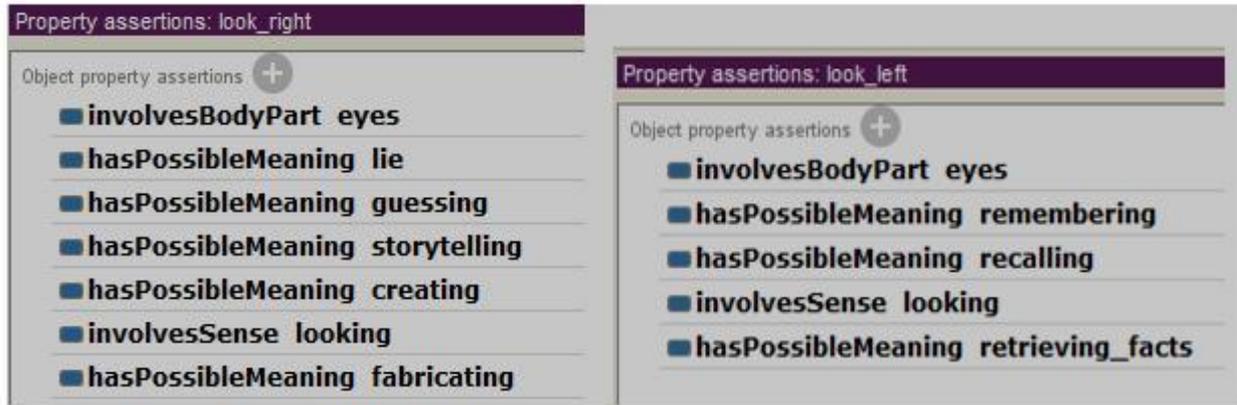


Fig. 8. Object property of 'look_right' and 'look_left' individuals.

```
SELECT ?individual
WHERE{ onto:look_right onto:hasPossibleMeaning ?individual}
```

Fig. 9. Eye signal meaning SPARQL query.

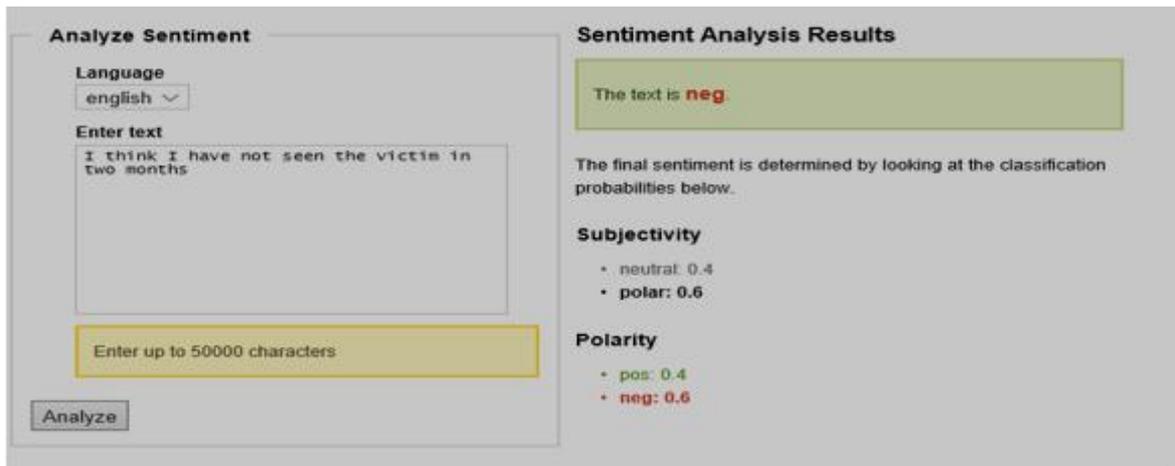


Fig. 10. Negative sentiment analysis result of statement.

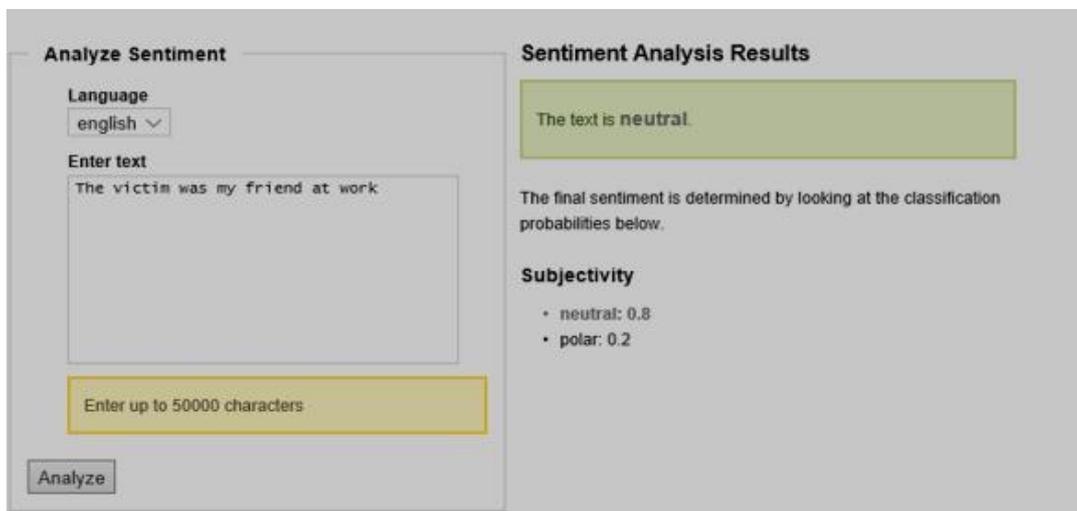


Fig. 11. Neutral sentiment analysis result of statement.

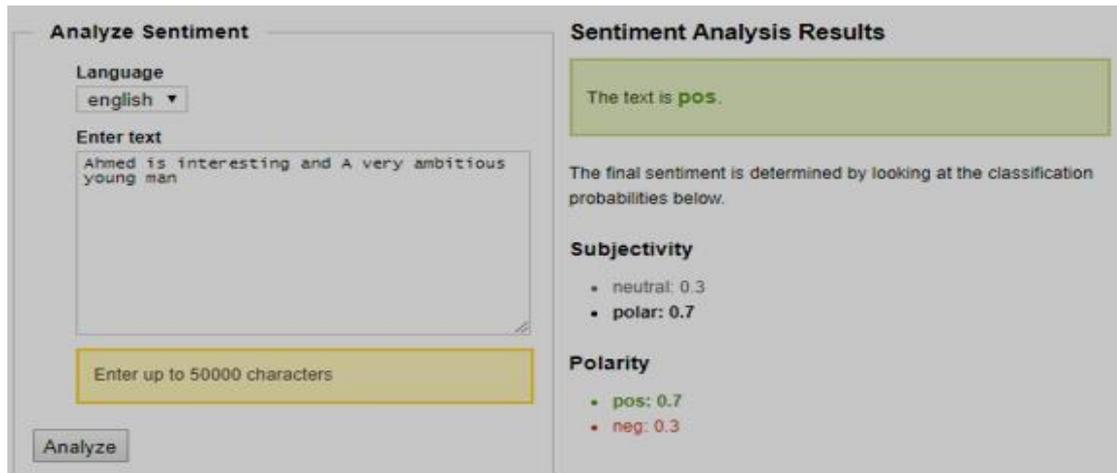


Fig. 12. Positive sentiment analysis result of statement.

4.3. Final Report Phase

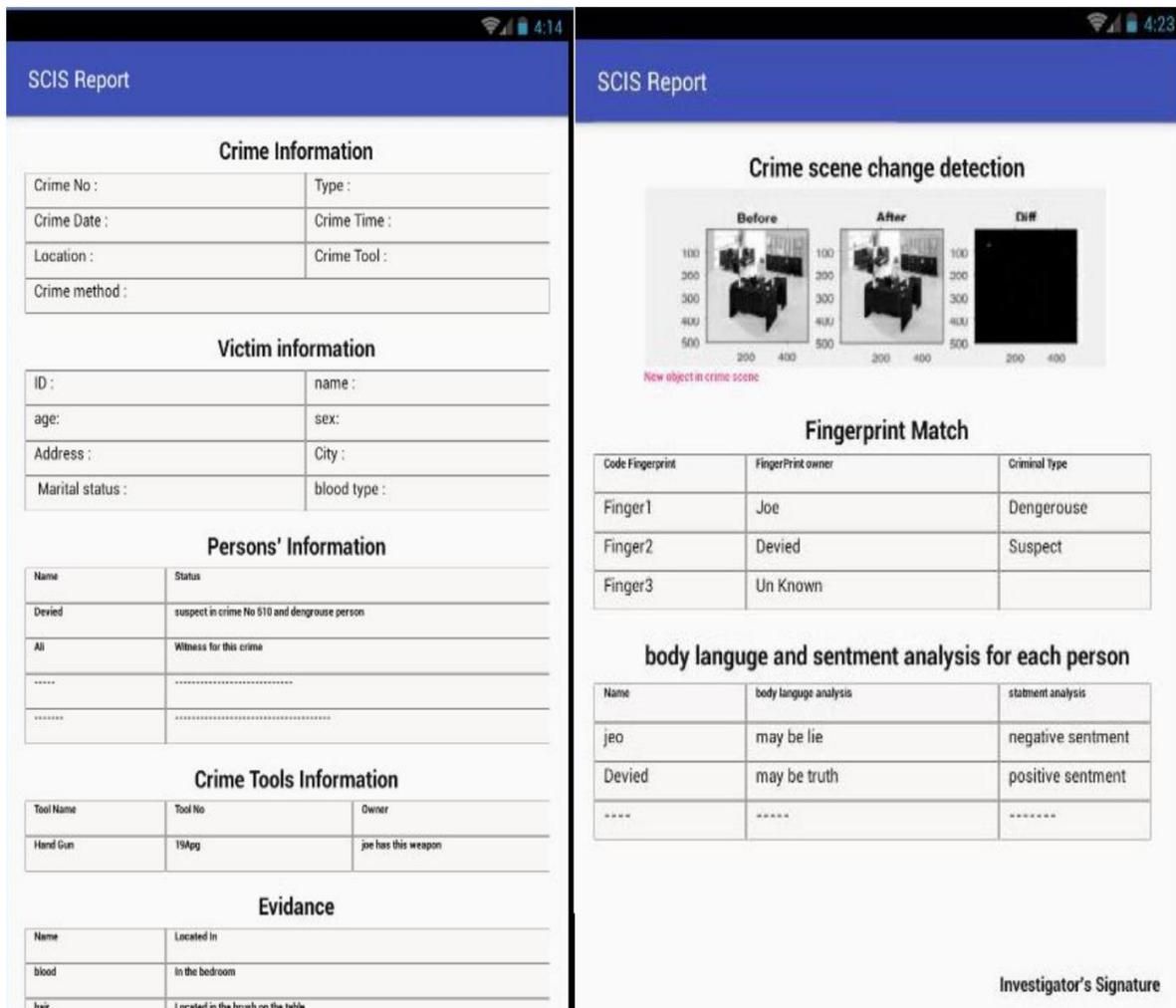


Fig. 13. The system output (SCIS Report).

The investigation report can be one of the most valuable sources of knowledge which SICS can extracted and merged from evidences. Writing an investigation report is one of the most hardest tasks an investigator undertakes. It is important that the report is detailed, simple and accurate, but does not include unnecessary

details. The final report is designed as mobile application using Android Studio 2.3.2. Fig. 13 presents snapshot of the system output (SCIS Report).

5. Conclusion

In this article we proposed SCIS (semantic crime investigation system). SCIS is an intelligent multi-phases crime investigation system based on Ontology re-engineering for representing main concepts of police criminal investigation and the relations among the concepts in the domain. Using Ontology leads to improve the quality of system design.

Fusion the direct and indirect evidences ameliorate the SCIS decision. Where, the direct evidences collected from crime scene or Civil Registry database and the indirect evidences collected from body language as (eye movement) and statement sentiment classification of speaker during the investigation. The investigator can access to the system via mobile application. Finally, the proposed system is general, scalable, mobile, fast and semantic to support the investigator to take a certainty decision.

In the future, we can merge the Universal Ontology with Gene Ontology (GO) directly for more certainty fast knowledge. This addition in knowledge based contributes for enhancement extract biometric features and relations quickly. Also, we can consider SCIS as the first step for solving International crime, when we using Big data concepts with our proposed system.

References

- [1] Yamuna, S., & Bhuvaneshwari, N. S. (2012). Data mining techniques to analyze and predict crimes. *The International Journal of Engineering and Science*, 1(2), 243-247.
- [2] Duan, L. (2011). Cooperative expert system based on extended intelligence: A new application of criminal investigation. *Advances in Mathematical and Computational Methods*, 1(2), 43-52.
- [3] Dzemydiene, D., & Kazemikaitiene, E. (2005). Ontology-based decision support system for crime investigation processes. *Information Systems Development*, 427-438.
- [4] Sir, M., Bradac, Z., & Fiedler, P. (2015). Ontology versus database. *IFAC-Papers Online*, 48(4), 220-225.
- [5] Nattawuttisit, S., & Usanavasin, S. (2012). Ontology mapping based on machine learner. *International Journal of Research in Engineering and Technology (IJRET)*, 1(2), 335-339.
- [6] Laurini, R., & Kazar, O. (2017). Geographic ontologies: Survey and challenges. *Meta-Carto-Semiotics*, 9(1), 1-13.
- [7] Chang, C., & Hua, K. (2008). Applying case-based reasoning and expert systems to coastal patrol crime investigation in Taiwan. *Intelligence and Security Informatics*, 161-170.
- [8] Keyvanpour, M. R., Javideh, M., & Ebrahimi, M. R. (2011). Detecting and investigating crime by means of data mining: A general crime matching framework. *Procedia Computer Science*, 3, 872-880.
- [9] Onnoom, B., Chiewchanwattana, S., Sunat, K., & Wichienit, N. (2014). An ontology design for recommendation filling-in for a crime scene investigation report of the forensic science police center 4 thailand. *Advanced Approaches to Intelligent Information and Database Systems*, 551, 139-147.
- [10] Falaye, A. A., Adama, N. V., & Agemerien, F. P. (2013). Design and implementation of crime investigation system using biometric approach (nigerian police force). *The Pacific Journal of Science and Technology*, 14(2), 242-253.
- [11] Alzaabi, M., Taha, K., & Martin, T. A. (2015). Cisri: A crime investigation system using the relative importance of information spreaders in networks depicting criminals communications. *IEEE Transactions on Information Forensics and Security*, 10(10), 2196-2211.
- [12] Elsayed, E., & Ghanam, M. (2016). SGEDSS: Semantic gene expression model for communication decision support system. *Journal of Computer Engineering*, 18, 46-52.

- [13] Protégé Platform. Retrieved from <http://protege.stanford.edu/>
- [14] Zoric, G, Smid, K., & Pandzic, I. S. (2007). Facial gestures: Taxonomy and application of non-verbal, non-emotional facial displays for embodied conversational agents. *Conversational Informatics*.
- [15] Leal, S., & Vrij, A. (2008). Blinking during and after lying. *Journal of Nonverbal Behavior*, 32(4), 187-194.
- [16] Despotakis, D., Thakker, D., Lau, L., & Dimitrova, V. (2011). Capturing the semantics of individual viewpoints on social signals in interpersonal communication. *Semantic Web Journal*.
- [17] Song, X., Liang, J., Hu, C., Kim, C., Lee, S., Kim, Y., & Kamdar, M. (2016). Sentiment classification: A topic sequence-based approach. *JCP*, 11(1), 1-9.
- [18] Python NLTK text classification demo. Retrieved from <http://text-processing.com/demo/sentiment/>



Eman K. Elsayed received the Ph.D in computer science 2005 at Alazhar University, received the master of computer science at Cairo University in 1999, received the bachelor of science, mathematics and computer science at Department, Cairo University 1994. She published thirty three articles until now in data mining, ontology engineering, e-learning, image processing, speech processing, machine learning and software engineering. She published two books in formal methods and event B on Amazon database.

She is a member in Egyptian mathematical society and intelligent computer and information systems society.



Wafaa Ghonaim received her B.Sc degree with honors in computer science and pure mathematics from faculty of Science, Al-Azhar University, in 2005. She received her M.Sc and Ph.D in computer science from Faculty of Science, Al-Azhar University, Cairo, Egypt in 2009, 2013. She is working as a lecture at Faculty of Science, Al-Azhar University. Her current research interest is in the area of computational intelligence based on rough set theory, swarm optimization, image processing, and image retrieval. In addition, she is

interested in the area of evolutionary computation in cryptanalysis.



Eman Zear received her B.Sc degree with very well in computer science and pure mathematics from the faculty of science, Al-Azhar University in 2012. She is working as a developer; her current research interest is in the area of artificial intelligence, programming, database.