

## SICH Project Final Conference

The SICH project was created to support the Police Forces, and all the other European organizations active in facing and preventing online crimes, in facilitating and speeding up the activities of identification and removal of illegal text content from all the information sources available online.

Online monitoring and analysis of illegal content had always paid much attention to multimedia content processing such as images, videos, etc., giving lower attention to text files. This is mainly due to the fact that texts (communications, conversations, news, websites, blogs) are mostly **unstructured data** that rarely have an organized architecture like, for example, the one we can find on a database.

As a consequence, the analysis and processing activities of unstructured information are sometimes inefficient, laborious and time-consuming. Moreover, much of the useful information risk to be ignored by analysts due to the large amount of data and due to its continuous change.

However textual resources represent a substantial part of web resources and the information gleaned from them, through the use of semantic technology, are fundamental to achieve an effective monitoring and to prevent and face crime and illegal content online. At the moment, there are no tools to deal with this problem or available in the public domain.

The Police Forces, research institutes, media and international organizations active in the research and discovery of illegal online content can face great difficulties in collecting and managing the large amount of content related to a specific domain.

The tools able to support analysts in the identification, selection and analysis of illegal online content, can improve the understanding of the phenomenon and recognition of illegal content hidden in textual information on the web.

SICH intends to create an **innovative model** capable of supporting analysts in the investigation activities and in the analysis of illegal content on Internet, combining the criminological approach to criminal phenomena, belonging to the three focus areas of the project (**Xenophobia/Racism, Illegal Online Gambling and NPS - New Psychoactive Substances**), and applying semantic technology.

This approach allows end-users to achieve better and faster understanding of crime on the Internet, to get instantaneous and targeted answers as part of the investigation, to analyze and identify potentially dangerous trend and to support cooperation between domain experts, law enforcement and private entities.



### Logistica / Logistic

HOTEL DIANA - Roof Garden Via Principe Amedeo, 4 00185 ROMA - (ITALY)

PHONE ++39.06.478681 FAX ++39.06.486998

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

[info@hoteldianaroma.com](mailto:info@hoteldianaroma.com)

### Contatti / Contacts

Alessandro Capone | Expert System S.p.A.

Mobile +39 3346766032

[acapone@expertsystem.com](mailto:acapone@expertsystem.com)

Claudia Gallo | Expert System S.p.A.

[cgallo@expertsystem.com](mailto:cgallo@expertsystem.com)

Alessio Mezzetti | Expert System S.p.A.

[amezzetti@expertsystem.com](mailto:amezzetti@expertsystem.com)



Co-funded by the Prevention of and Fight against  
Crime Programme of the European Union

## PROGETTO SICH Semantic Illegal Content Hunter



# 20<sup>2015</sup> CONFERENZA FINALE NOVEMBRE

ROMA Hotel Diana  
Via Principe Amedeo 4

NOVEMBER 20th 2015  
Final Conference – SICH  
Semantic Illegal Content Hunter  
ISEC Project



Co-funded by the Prevention of and Fight against  
Crime Programme of the European Union

# Conferenza finale del progetto SICH

Il progetto SICH nasce per supportare le Forze di Polizia e le altre organizzazioni europee attive nella lotta e nella prevenzione dei crimini online, facilitando e velocizzando le attività di identificazione e rimozione dei contenuti testuali illegali provenienti da tutte le fonti di informazione disponibili on-line.

Da sempre le attività di monitoraggio e analisi dei contenuti illegali online hanno dedicato molta attenzione al trattamento dei file multimediali quali immagini, video, ecc., riservando minor interesse per i file testuali. Ciò è dovuto principalmente al fatto che i testi (comunicazioni, conversazioni, notizie, siti web, blog) sono per lo più **dati non strutturati** che raramente hanno una forma organizzata, simile, cioè, a quella di un database. Di conseguenza, le attività di analisi ed elaborazione delle informazioni non strutturate risultano talvolta inefficienti, faticose e time-consuming. Inoltre, gran parte delle informazioni utili rischia di essere ignorata dagli analisti a causa della grande quantità e del continuo cambiamento.

Tuttavia le risorse testuali rappresentano una parte consistente delle risorse web e le informazioni estrapolate da esse, tramite l'utilizzo di tecnologia semantica, risultano fondamentali per effettuare un monitoraggio efficace, per prevenire e combattere la criminalità ed i contenuti illegali online. Al momento, non esistono strumenti in grado di affrontare questo problema o disponibili nel pubblico dominio.

Le forze di Polizia, gli istituti di ricerca, i media e le organizzazioni internazionali attive nella ricerca e scoperta dei contenuti illegali online possono incontrare grandi difficoltà nella raccolta e gestione della grande mole di contenuti legati ad uno specifico dominio. Gli strumenti in grado di supportare gli analisti nell'identificazione, selezione e analisi dei contenuti illegali online possono migliorare la comprensione del fenomeno e il riconoscimento dei contenuti illegali nascosti nelle informazioni testuali presenti sul Web.

Combinando l'approccio criminologico ai fenomeni criminali, appartenenti alle tre aree di interesse del progetto (**Xenofobia/Razzismo, Gioco d'azzardo illegale online e Nuove Sostanze Psicoattive - NPS**), e applicando la tecnologia semantica, SICH intende creare un **modello innovativo** capace di supportare gli analisti nelle attività di investigazione e analisi dei contenuti illegali in rete.

Questo tipo di approccio consente agli utilizzatori finali di ottenere una migliore e più rapida comprensione dei fenomeni criminali su Internet, risposte immediate e mirate nell'ambito delle indagini, analisi ed identificazione di trend potenzialmente pericolosi e supporto alla cooperazione fra esperti di dominio, forze dell'ordine e soggetti privati.

09.00 - 09.30  
Registrazione dei partecipanti  
*Welcome and Registration*

09.30 - 09.45  
Il progetto SICH  
*The SICH Project*

Crimine 2.0: Internet come strumento a supporto del crimine. I casi delle Nuove Sostanze Psicoattive (NPS) e dei crimini d'odio per finalità di razzismo e xenofobia (online hate speech)  
*Crime 2.0: Internet as crime facilitator. The case of Novel Psychoactive Substances (NPS) and online hate*

09.45 - 10.15  
Cybercrime: vecchi e nuovi scenari  
*Cybercrime scenarios*

10.15 - 10.45  
La crescita e la diffusione delle Nuove Sostanze Psicoattive  
*The emerging of Novel Psychoactive Substances*

10.45 - 11.15  
L'evoluzione dell'Online Hate Speech  
*The evolution of Online Hate Speech*

11.15 - 11.45  
Coffee Break

11.45 - 12.25  
Il contrasto ai crimini online - case histories in ambito comunitario  
*The fight against Cybercrime - case histories from European experts*

12.25 - 12.40  
Sessione Aperta: Q&A  
*Open Discussion: Q&A*

12.40 - 13.40  
Lunch

13.40 - 14.10  
Realizzazione di un motore di ricerca ottimizzato per la ricerca di contenuti illegali  
*A search engine optimized for discovering illegal contents*

14.10 - 15.10  
SICH FOCUS: presentazione ed utilizzo della piattaforma SICH per l'analisi del fenomeno NPS e dell'online hate speech  
*SICH FOCUS: Using SICH PLATFORM to analyse NPS and Online Hate Speech over the Internet*

15.10 - 15.25  
Sessione aperta: Q&A  
*Open Discussion: Q&A*

15.25 - 15.55  
Coffee break

15.55 - 16.40  
Uno sguardo verso il futuro - l'eredità del progetto SICH  
*A look towards the future - SICH's legacy*

Conclusioni  
*Conclusions*

16.40 - 17.30  
Sessioni Pratiche di utilizzo della piattaforma  
*Practical Sessions*

