

Workshop e Tavola Rotonda

La scienza dei dati e la formazione della prova nelle aule giudiziarie

Palermo, 01 Luglio 2023
Palazzo Steri Sala Magna
Università degli Studi di Palermo
Piazza Marina 61 90133 Palermo

09:00 09:30	Saluti istituzionali Prof. Massimo Midiri, Magnifico Rettore UNIPA A. Elio Mineo, Direttore DSEAS G. Massimo Palma, Direttore DiFC
------------------------------	--

SESSION I: chairman Salvatore Micciché – DiFC

09:30 09:50	Maurizio de Lucia Procuratore della Repubblica presso il Tribunale di Palermo	Presentazione della Convenzione tra Procura della Repubblica e i dipartimenti universitari DSEAS e DiFC.
09:50 10:10	Francesco Petruzzella Procura della Repubblica di Palermo	Le infrastrutture dei dati nelle procure italiane: stato dell'arte, opportunità e rischi
10:10 10:30	Ten. Col. GdF Paolo PETTINE Nucleo di polizia economico-finanziaria di Palermo	Digital forensics & digital investigation

10:30 11:00	Coffee break
------------------------------	---------------------

SESSION II: chairman Michele Tumminello – DSEAS

11:00 11:15	Salvatore Micciché UNIPA DiFC	Tecniche di <i>data mining</i> basate su reti complesse: dalla conoscenza accademica alla formulazione di ipotesi investigative empiricamente fondate.
11:15 11:30	Michele Tumminello UNIPA DSEAS	Rilevazione di truffe assicurative attraverso un'analisi di rete dell'Archivio Integrato Antifrode.
11:30 11:45	Giosuè Lo Bosco UNIPA DMI	Data driven artificial intelligence per l'estrazione di conoscenza da immagini e testi.

11:45 12:45	Tavola Rotonda Moderatore: Andrea Consiglio UNIPA DSEAS
------------------------------	---

12:45 13:00	Saluti Finali
------------------------------	----------------------

Il workshop è aperto a Magistrati, Forze dell'Ordine e Ricercatori.

In questo workshop si intende mettere a confronto le esperienze di ricercatori che lavorano nel campo della Data Science con quelle di chi professionalmente si occupa di indagini giudiziarie: rappresentanti delle forze dell'ordine e magistrati.

L'obiettivo del workshop è duplice. Da una parte si vuole portare all'attenzione degli operatori professionalmente deputati allo svolgimento delle indagini quali possono essere le opportunità che le metodologie legate al *data mining*, alla scienza delle reti complesse, alle tecniche di intelligenza artificiale come *machine learning*, *deep learning* e *natural language processing* possono costituire per il loro lavoro. Opportunità che possono riguardare tanto aspetti specifici quanto la formulazione di scenari investigativi da approfondire con tecniche di indagine più tradizionali.

D'altra parte, si vuole portare all'attenzione del mondo accademico quali siano le esigenze degli operatori di giustizia, sia dal punto di vista degli strumenti operativi (forze dell'ordine), sia dal punto di vista della loro validazione e certificazione in relazione al loro uso nelle aule giudiziarie (magistrati).

Per informazioni rivolgersi a:

michele.tumminello@unipa.it
salvatore.micciche@unipa.it

APPROFONDIMENTO

Il progetto culturale che sta alla base di questo workshop si inquadra nell'ambito della *Legal Analytics* che è un moderno strumento di analisi che applica al diritto tecnologie per l'estrazione di informazioni, da grandi quantità di dati, attraverso metodi automatici o semi-automatici e l'utilizzo scientifico e operativo delle stesse;

Obiettivo a lungo termine di questo progetto è la creazione di un protocollo di lavoro condiviso che preveda, ad esempio, l'analisi testuale e il riconoscimento di immagini, con tecniche di *Artificial Intelligence* (*machine learning* e *natural language processing*), partendo dall'analisi di grandi quantità di dati documentali, al fine di individuare relazioni statisticamente significative tra gli elementi presenti, passivamente o attivamente, nel contesto giudiziario, sociale ed economico a cui fanno riferimento i testi e le immagini prese in considerazione.

Il materiale da sottoporre ad analisi può provenire da fonti ad accesso riservato (giudiziarie e investigative) e da fonti aperte (media, open data, big data istituzionali), ciascuna classificabile per livello di affidabilità e attendibilità.

Tra gli aspetti che si intendono approfondire nella discussione citiamo:

- la validazione statistica di interrelazioni tra elementi (soggetti, attività, ecc.) estratti da testi ed immagini con tecniche di AI
- lo sviluppo di algoritmi di AI per problemi di ottimizzazione e apprendimento su reti complesse di parole e documenti.
- l'integrazione tra metodi di fisica statistica, statistica multivariate e teoria delle reti per determinare modelli minimali di interrelazioni significative;
- La descrizione sistematica dei modelli di intelligenza artificiale utilizzati e sviluppati in modo tale che siano certificabili in relazione ad un eventuale uso in dispute legali;
- l'ottimizzazione degli algoritmi al fine di arrivare (a regime) a soluzioni implementabili in *real-time* da parte degli *stakeholders*;
- classificazione *a posteriori* dell'affidabilità delle fonti.