Tutta l'intelligenza (artificiale) nelle mani della polizia

Riccardo Coluccini @ORARiccardo

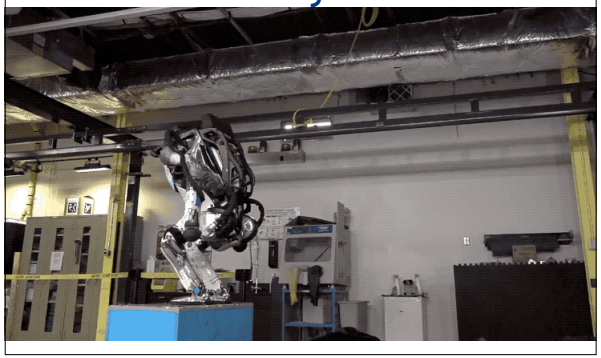
E-PRIVACY XXIII

8th-9th June 2018

HERMES CENTER FOR TRANSPARENCY AND DIGITAL HUMAN RIGHTS

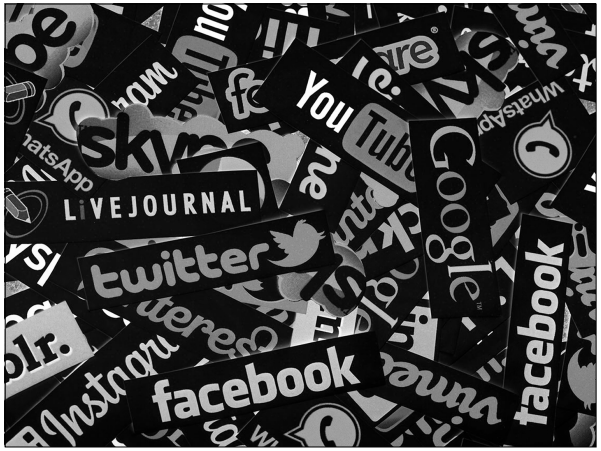
AI

AI Reality Check



The Switch

Amazon is selling facial recognition to law enforcement — for a fistful of dollars



Voice Recognition (CRAIM)

2.1 Software e Apparati

L'oggetto della fornitura è rappresentato dal complesso degli apparati, dei servizi e delle attività come descritti nel presente capitolato tecnico. Si elencano in maniera sintetica i macro elementi del progetto CRAIM che sono oggetto di fornitura:

- **Antenna Parabolica:** sistema fisico per la ricezione di trasmissioni televisive da satelliti da così detti "canali in chiaro".
- **Sistema di registrazione** sistema comprensivo di apparati hardware, per la registrazione e la transcodifica di segnali DVB-T, DVB-T2, DVB-C DVB-S e DVB-S2.
- **Broadcasting Monitoring** componente logica per l'acquisizione (streaming/registrazione) e la trascrizione automatica dei contenuti audio dei canali radiotelevisivi prescelti e la fruizione dei contenuti audio/video sincronizzati con il testo trascritto.
- **Media Monitoring** componente logica per la trascrizione automatica di fonti audio/video presenti su web, mediante l'ausilio di strumenti di speech to text multilingua in grado di effettuare crawling e speaker identification (identificazione automatica e ricerca di determinate impronte vocali).
- **Knowledge Management Semantico-ontologico**: componente logica per l'indicizzazione dei contenuti trascritti e la ricerca dei temi di interesse, finalizzata alla più efficace ricerca delle informazioni acquisite in modalità multicanale.
- **Database impronte vocali**: componente logica che si occupa di archiviare dati "voce" con dati anagrafici e li rende disponibili al motore di ricerca e confronto per individuare quelli rispondenti a soggetti noti.

CRAIM

- Almawave

iridevoice

Iride Voice traduce attraverso sistemi di riconoscimento vocale, prodotti dalla controllata Pervoice, una conversazione telefonica in testo e, grazie all'interpretazione semantico ontologica del linguaggio naturale, consente di identificare, comprendere e classificare in modo semplice, veloce ed efficace le esigenze dei clienti, rintracciando ed estraendo i concetti di interesse.

La soluzione dispone di advanced analytics e knowledge management system, realizzati per disporre di dati utili per ridisegnare i processi di churn prevention, marketing insight, claims, training.

CRAIM

- FOIA request to obtain technical offers

Di conseguenza, l'accesso richiesto non viene consentito in relazione alle offerte tecniche presentate dalla R.T.I. "CEDAT 85 S.r.l./I.B.M. Italia S.p.A." e dalla Società "ALMAWAVE S.p.A.".

Attesa la mancata ricezione di opposizione da parte della altri quattro operatori e valutata l'assenza di motivi ostativi per quanto attiene alla tutela degli interessi pubblici elencati nell'articolo 5 bis, comma 1, del D. Lgs.vo nr. 33/2013, l'istanza di accesso formulata dalla S.V. è invece accolta, mediante ostensione di copia delle offerte tecniche presentate rispettivamente dalla Società "Telecom Italia S.p.A.", dalla Società "RCS S.p.A.", dalla Società "VITROCISET S.p.A." e dal R.T.I. "HP ITALIANA S.r.l./BUSINESS-E S.p.A.".

IL DIRIGENTE

CRAIM

3.4.1 ***Componente ValueCore DeepView***

L'intera soluzione è stata progetta utilizzando un modello di processo basato sula metodologia delle Four D, proveniente da varie declinazioni di ricerca in ambito OSINT (Open Source INTelligence) descritta nella figura seguente:

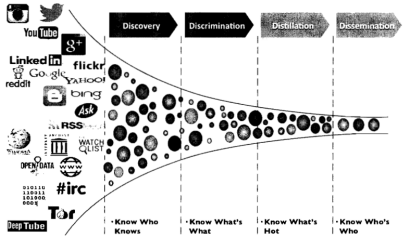


Figura 5: Metodologia delle Four D

FourD Metodology – Fase 1 - Discovery

CRAIM

- A solution for voice fingerprints and database

Features

A complete public security solution

The full range of voice biometric technology from Nuance and Agnitio.

Nuance Forensics

Based on Nuance's innovative industry-leading algorithms, Nuance Forensics provides forensic examiners and investigators the ability to accurately match an individual's identity with content captured through any type of audio channel. It compares an audio file to a speaker profile and a reference population to statistically assess a match. Nuance Forensics is a complete set of tools to develop reference populations, build speaker profiles, and construct a library of certified speaker samples.

Nuance Identifier

Through voice biometric analysis of audio files, Nuance Identifier can identify known individuals of interest, enabling government officials to match recorded conversations to targets. Using industry-leading text-independent voice biometric algorithms, Nuance Identifier can help officials find the needle-in-a-haystack by performing 200,000 voice biometric comparisons per second per CPU core. Large scale searches on millions of audio files can be performed within minutes.

Agnitio Public Security products

Agnitio has been a leader in the Public Security market for over a decade. Agnitio Voice ID is used by Government organizations to prevent crime, identify criminals and provide evidence in court. Agnitio has an extensive customer base including leading police, intelligence, military, and other government organizations in over 40 countries. Agnitio products for public safety include BATVOX, ASIS, BS3 and SIFT.

CRAIM

- Nuance identifier

Deep neural networks-based voice biometrics

Nuance voice biometric algorithms have been used to protect security-critical applications since 2001. In 2015, Nuance released the industry's first voice biometric algorithms powered by deep neural networks (DNN), a computer learning technology that enables a quantum leap in performance. Nuance Identifier is embedding the third generation of its DNN-based voice biometric algorithms, setting a new industry benchmark in voice biometric performance.

Face Recognition (SARI)

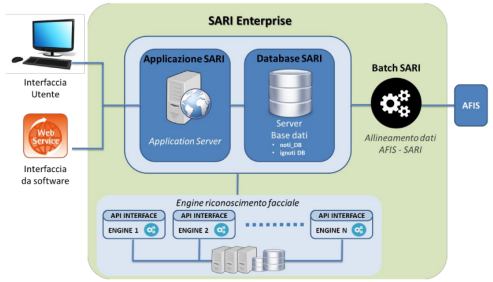


Figura 1-Architettura logica Sistema Enterprise

SARI

- Police considers SARI as an upgrade to AFIS

"SARI rappresenta l'evoluzione tecnologica di un sistema che già abbiamo," spiega Caterina Angelicchio, direttrice generale della prima sezione della seconda divisione del servizio di Polizia Scientifica.

"Il precedente sistema consentiva agli organi di investigazione di fare delle ricerche per anagrafica, e secondo particolari caratteristiche all'interno della banca dati AFIS (Automated Fingerprint Identification System), in cui sono contenuti tutti i dati dei cartellini fotosegnaletici delle persone."

SARI

- Specs

Dimensionamento dello Scenario Sistema Enterprise ambiente di produzione	
Numero di immagini da dover gestire	10.000.000
Numero di utenti simultaneamente loggati	100
Numero di ricerche contemporanee	10
Tempi di risposta per il riconoscimento facciale	< 15 secondi

Dimensionamento dello Scenario Real-Time	
Numero di immagini della watch-list	10.000
Numero di fotogrammi al secondo per ogni telecamera	15
Numero di telecamere CCTV	10
Numero soggetti per fotogramma	6
Risoluzione telecamera	5 MPixel
Tempo di generazione dell'alert	< 2 secondi

SARI

- Parsec 3.26

Il sistema S.A.R.I. realizzato da Parsec 3.26 per il Ministero dell'interno, supera positivamente le verifiche di collaudo

Sistema S.A.R.I. - si procede verso la fase di avviamento su tutto il territorio Nazionale.

Parsec è soddisfatta del risultato ottenuto e della qualità del lavoro svolto nel corso dell'ultimo anno, che ha permesso di superare positivamente le verifiche di collaudo del sistema S.A.R.I. Adesso si procederà con la fase di avviamento, che prevede dapprima, l'erogazione di corsi formativi alle Forze di Polizia su tutto il territorio Nazionale. Si partirà dalla Capitale, mentre il roll-out in tutte le altre regioni verrà avviato nei prossimi mesi.

SARI

- Face recognition algorithm discovered via FOIA

SARI

- NIST 2014

Accuracy across commercial providers: Recognition accuracy is very strongly dependent on the algorithm and, more specifically, on the developer of the algorithm. Recognition error rates in a particular scenario range from a few percent up to beyond fifty percent. Among the most accurate developers, the rank one miss rates for recognition in a population size of 1.6 million are 4.1% (NEC), 9.1% (Morpho), 10.7% (Toshiba), 13.6% (Cognitec), 17.2% (3M) and 20.5% (Neurotechnology). For webcam images, this sequence is 11.3% (NEC), 23.7% (Toshiba), 29.8% (Morpho), 36.4% (3M), 57.6% (Cognitec) and 66.9% (Neurotechnology). While results for up to six algorithms from each developer are reported here, the intra-provider accuracy variations are usually smaller than the inter-provider variations. That said, some developers submitted different, less accurate but computationally lightweight algorithms.

SARI

- Reco algorithm

Se la foto non è idonea per l'uso che se ne vuole fare, può essere eliminata cliccando il pulsante "***ANNULLA***". L'applicativo visualizzerà così, nuovamente, la schermata mostrata in figura 1, in modo da poter acquisire una nuova foto dal soggetto. Se la foto invece è idonea, l'operatore clicca il pulsante "**INVIA FOTO**", facendo comparire una schermata come la seguente.

Figura 3 – Verifica di un soggetto sospetto – Schermata 2/6

L'operatore seleziona, mediante gli appositi controlli, il "***sesso***" e l' "***etnia***" del soggetto sospetto, e successivamente clicca il pulsante "**INVIA**". La foto, insieme ai dati imputati, viene inviata ai "*componenti di ricerca*" del sistema SARI Real-Time che la sottopongono ad elaborazione. Se le "componenti di ricerca" rilevano, nella watch-list, foto di soggetti attenzionati "simili" a quella del sospetto, ricevuta in input, le restituisce all'applicativo mobile, sotto forma di elenco. Compare una schermata come la seguente.

Figura 4 – Verifica di un soggetto sospetto – Schermata 3/6

SARI

- Reco 3.26

SARI

- FOIA request to obtain statistics on SARI

Documenti e informazioni	1) Data di inizio e fine collaudo del sistema SARI. Come indicato online dall'azienda Parsec 3.26, la fase di collaudo si è già conclusa: https://is.gd/dL9sDJ 2) Tutti i documenti relativi all'esito del collaudo. 3) Eventuali documenti relativi ad analisi preliminari sull'impatto del sistema SARI sulla privacy dei cittadini. Inclusi documenti ricevuti o inviati al Garante per la protezione dei dati personali. 4) Dettaglio del numero di match positivi prodotti dal sistema durante il suo utilizzo. 5) Dettaglio del numero di falsi positivi prodotti dal sistema durante il suo utilizzo. Come riferimento, sarebbe opportuno ottenere un dettaglio simile a quello fornito dalla South Wales Police: https://www.south-wales.police.uk/en/advice/facial-recognition-technology/

Face Recognition

- False Positive figures by South Wales Police

Event	True Positive Alerts	False Positive Alerts
UCL	173	2,297
Elvis Festival	10	7
Op. Fulcrum	5	10
Joshua Fight	5	46
Wales vs Australia	6	42
Wales vs Georgia	1	2
Wales vs New Zealand	3	9
Wales vs South Africa	5	18
Kasabian	4	3

AI Risks & Challenges

- Protection of biometric data
- Consumer-oriented AI products repurposed for LEA
- Chilling effect
- Lack of accountability due to undisclosed statistics
- Data violence

Thank you!

- Contacts

  riccardo.coluccini@hermescenter.org

E-PRIVACY XXIII

8th-9th June 2018

HERMES CENTER FOR TRANSPARENCY AND DIGITAL HUMAN RIGHTS