



## ACCORDO DI COLLABORAZIONE

(ex art. 15 della Legge 241/1990)

Co.RE.COM. CALABRIA

UNIVERSITÀ DEGLI STUDI DI MESSINA

UNIVERSITÀ DEGLI STUDI "MEDITERRANEA" DI REGGIO CALABRIA

## PIANO DELLE ATTIVITÀ

<b>1.Mission del CO.RE.COM. CALABRIA.</b>	<b>3</b>
<b>1.1 Declinazione delle attività di competenza del CO.RE.COM. Calabria.</b>	<b>4</b>
<b>1.2 Implementazione di Progetti di Ricerca - Sperimentazione-Analisi su altre tipologie di Flussi Audio - Video Generate da altre fonti (RADIO-WEB-SOCIAL).</b>	<b>6</b>
<b>1.3 Aggiornamento – Formazione.</b>	<b>7</b>
<b>2. LINEE DI RICERCA DELLE UNIVERSITÀ.</b>	<b>7</b>
<b>2.1 Università degli Studi di Messina (UniME).</b>	<b>8</b>
<b>2.1.1 Dipartimento di Scienze Matematiche e Informatiche, Scienze Fisiche e Scienze della Terra (MIFT).</b>	<b>8</b>
<b>2.2. Università degli Studi “Mediterranea” di Reggio Calabria.</b>	<b>9</b>
<b>2.2.1 Dipartimento di Ingegneria Civile, Energia, Ambiente e Materiali (DICEAM).</b>	<b>9</b>
<b>3.SPERIMENTAZIONE OPERATIVA.</b>	<b>10</b>
<b>3.1 Il CO.RE.COM. Calabria come test-case reale.</b>	<b>11</b>
<b>3.1.1 Il Cloud computing come base dei Sistemi Innovativi.</b>	<b>11</b>
<b>3.1.2 Sperimentazioni legate all’utilizzo dei servizi Cloud per Audio -Video Acquisition.</b>	<b>13</b>
<b>3.1.3 Sperimentazioni legate all’utilizzo dei servizi Cloud per storage dei flussi audio-video.</b>	<b>13</b>
<b>3.1.4 Sperimentazioni legate all’utilizzo dei servizi Cloud per processing dei flussi audio-video e caratterizzazione.</b>	<b>14</b>
<b>3.1.5 Applicare un nuovo Approccio a Blocchi.</b>	<b>14</b>
<b>3.1.6 Sicurezza del sistema informativo.</b>	<b>15</b>
<b>3.1.7 Applicazione del Machine e Deep Learning.</b>	<b>15</b>
<b>3.1.8 Interagire con grosse moli di dati: Big Data.</b>	<b>17</b>
<b>3.1.9 Guardare dentro ai dati: Big Data Analytics.</b>	<b>18</b>
<b>3.1.10 Nuove soluzioni attraverso l’uso della Blockchain.</b>	<b>19</b>
<b>4. RISORSE UMANE IMPIEGATE NEL PROGETTO.</b>	<b>20</b>
<b>4.1 Risorse umane impiegate dalle Università.</b>	<b>20</b>
<b>4.2 Risorse umane impiegate dal Co.Re.Com Calabria.</b>	<b>21</b>
<b>4.3 Interazione gruppi di lavoro.</b>	<b>21</b>
<b>4.4 Gantt delle attività.</b>	<b>22</b>
<b>Design e setup di algoritmi e architetture di ML/DL per l’elaborazione dei flussi video</b>	<b>23</b>

<b>Descrizione Attività</b>	<b>24</b>
<b>Design e setup di algoritmi e architetture di ML/DL per la classificazione di sequenze video</b>	<b>24</b>
<b>Design, sviluppo e setup di moduli specifici di ML/DL per l'estrazione automatica delle informazioni d'interesse da flussi</b>	<b>24</b>
<b>Design e setup di algoritmi e architetture di ML/DL per il riconoscimento/rilevamento di possibili attacchi hacker</b>	<b>24</b>
<b>5. Dotazione Hardware- progetto sperimentazione.</b>	<b>25</b>
<b>5.1 Ricevitore segnale.</b>	<b>25</b>
<b>5.2 Unità di storage locale: mini cloud privato.</b>	<b>25</b>
<b>5.3 Unità di computazione remota: cloud pubblico Amazon</b>	<b>26</b>
<b>5.4 Unità di calcolo locale per Machine Learning</b>	<b>26</b>

# 1.Mission del CO.RE.COM. CALABRIA.

Il Comitato Regionale per le Comunicazioni (Co.Re.Com.) Calabria, svolge, oltre le attività previste dalla propria legge istitutiva (L.R. n.2/2001 ss.mm.ii.), anche quelle delegategli dall'Autorità per le Garanzie nelle Comunicazioni (AgCom) in virtù di delega, l'ultima sottoscritta il 19/21 dicembre 2017.

Tra le funzioni delegate, se, da un lato, è stato confermato all'art. 5, lettera f), la *“vigilanza sul rispetto delle norme in materia di esercizio dell'attività radiotelevisiva locale, mediante il monitoraggio delle trasmissioni dell'emittenza locale privata, nonché di quella della concessionaria pubblica, per l'ambito di diffusione regionale, con particolare riferimento agli obblighi in materia di programmazione, anche a tutela delle minoranze linguistiche e dei minori, pubblicità e contenuti radiotelevisivi previsti dal Tusmar, come integrato dai Regolamenti attuativi dell'Autorità”*, dall'altro, alla lettera a), sempre del citato articolo, gli è stato attribuito ex novo quella inerente alla *“tutela e garanzia dell'utenza, con particolare riferimento ai minori, attraverso iniziative di studio, analisi ed educazione all'utilizzo dei media tradizionali e dei nuovi media nel rispetto degli indirizzi stabiliti dall'Autorità e dalle Istituzioni competenti in materia, anche in attuazione di protocolli d'intesa per la migliore tutela decentrata dei minori sul territorio nazionale”*.

Si è riscontrata, quindi, in ambito di tutela e garanzia dell'utenza, con particolare riferimento ai minori, una diversa competenza, dall'estensione della vigilanza oltre i consueti palinsesti radio - televisivi locali.

Inoltre, l'AgCom ha predisposto un compendio di procedure operative per l'attività di monitoraggio e vigilanza sulle trasmissioni radiotelevisive in ambito locale, ad uso esclusivo dei Co.Re.Com. delegati, con cui ha definito le modalità esecutive del monitoraggio.

Nel contempo, ha fornito indicazioni in ordine alle procedure e agli indirizzi applicativi che regolano l'attività delegata, a cui i Co.Re.Com. sono stati invitati ad adeguarsi, allo scopo di assicurare una uniforme applicazione sul territorio nazionale della normativa richiamata, anche nell'intento di instaurare con i servizi di media audiovisivi locali un rapporto di collaborazione, mediante la promozione di una televisione di qualità.

Infine, l'Agcom ha - anche di recente - ribadito quanto aveva già raccomandato ai Co.Re.Com. delegati, sin dal conferimento delle prime deleghe in materia di vigilanza, ossia di dotarsi, in via principale, di sistemi di registrazione propri, per i vantaggi che tale soluzione offre, sia in ragione dei costi connessi, ritenuti contenuti e ascrivibili alle spese che giustificano il contributo da essa erogato, sia, soprattutto, in ragione dei vantaggi operativi derivanti dalla disponibilità del registrato sul territorio in caso di segnalazione, nonché per ovviare alle problematiche che possono insorgere, nel caso in cui gli operatori non collaborino, rifiutando di mettere a disposizione i dati registrati.

Stesso approccio, sempre verso la gestione in proprio, è stato fortemente sostenuto per quanto riguarda l'analisi e la classificazione del palinsesto televisivo, in cui è

fondamentale non trascurare di avere molto chiari i criteri di valutazione e determinazione delle fattispecie, per identificare le possibili violazioni ed organizzare il lavoro di analisi.

Ciò posto, il Co.Re.Com. Calabria ha colto queste raccomandazioni sin dalle prime deleghe (operative a decorrere dall'anno 2010), così che ha svolto entrambe le funzioni "in house"; vale a dire con personale dipendente appositamente autoformato, dotandosi, altresì, di propri sistemi di registrazione ed analisi.

Con l'ultima Convenzione, sulla scorta delle novità introdotte all'art. 5, considerate, parimenti, le mutate esigenze, inerenti la necessità di avviare iniziative di studio, analisi ed educazione dell'utenza, con particolare riguardo ai minori, all'utilizzo dei media tradizionali, nonché a quelli nuovi (web Tv, web radio, piattaforme social), per il Co.Re.Com. si sono aperte, accanto alle attività istituzionali già svolte negli anni scorsi, nuove prospettive di approfondimento con annessi nuovi modelli di analisi.

Su questa scia, tra i tipi di approccio che si vorrebbero implementare, vi è, ad esempio, quello concernente la sperimentazione di un sistema di c.d. "*Sentiment analysis*" o di "*Opinion mining*", inteso come il modo a cui ci si riferisce per indicare l'uso dell'elaborazione del linguaggio naturale, dell'analisi testuale e della linguistica computazionale per identificare ed estrarre informazioni soggettive da fonti diverse (come Tv e radio locali, web, social).

## **1.1 Declinazione delle attività di competenza del CO.RE.COM. Calabria.**

A) In linea e nel rispetto delle normative vigenti e delle direttive AgCom in materia di vigilanza, il Co.Re.Com. svolge l'attività di Acquisizione flussi audio-video marchi televisivi locali – Elaborazione – Analisi - Archiviazione dei dati - Presentazione dei dati.

- 1) Per acquisizione, si intende: acquisizione e storage dei flussi audio e video con trasmissione in digitale terrestre;
- 2) per elaborazione, si intende: elaborazione manuale del flusso audio-video;
- 3) per analisi, si intende: puntuale classificazione dell'emesso televisivo in base alle normative vigenti in materia di monitoraggio televisivo;
- 4) per archiviazione dei dati, si intende: repository dei dati su cloud locale ed estrazione dei dati in base alle necessità;
- 5) per presentazione dei dati, si intende:
  - a) generazione automatica di report secondo schemi AgCom;
  - b) elaborazione manuale di grafici/tabelle secondo necessità Co.Re.Com.

Le fasi di elaborazione ed archiviazione dei dati, finora condotte, vengono riassunte nella tabella di seguito riportata, nel rispetto delle macro e micro categorie (per i programmi, argomenti e soggetti politici/istituzionali/non istituzionali riportando

l'indicazione dell'ora di inizio e di fine) previste nella classificazione inserita nel manuale Agcom.

<b>Tipologia di dati da elaborare-analizzare ed archiviare</b>	<b>Macro Area di Monitoraggio</b>	<b>Reportistica</b>
<p>In relazione ai singoli argomenti e soggetti politici e/o istituzionali e/o sociali:</p> <ul style="list-style-type: none"> <li>Ø Tempo di notizia</li> <li>Ø Tempo di parola</li> <li>Ø Tempo di antenna</li> <li>Ø Tempo di Argomento (secondo macro e micro categorie Agcom)</li> </ul>	<p>Pluralismo Sociale e Politico</p>	<p>Generazione automatica di report secondo schemi Agcom (sia in formato PDF che Excel). Generazione automatica di Grafici</p>
<ul style="list-style-type: none"> <li>Ø Tempo di trasmissione e tipologia di programma</li> <li>Ø Individuazione, classificazione e report delle violazioni</li> </ul>	<p>Obblighi di programmazione</p>	<p>Generazione automatica di report secondo schemi Agcom (sia in formato PDF che Excel). Generazione automatica di Grafici</p>
<ul style="list-style-type: none"> <li>Ø Affollamenti e/o sforamenti (orari e giornalieri con calcolo automatico e detrazione dei frame neri).</li> <li>Ø Individuazione posizionamento degli eventi pubblicitari e dei contenuti</li> <li>Ø Individuazione, classificazione delle violazioni rilevate</li> </ul>	<p>Pubblicità</p>	<p>Generazione automatica di report secondo schemi Agcom (sia in formato PDF che Excel). Generazione automatica di Grafici</p>
<ul style="list-style-type: none"> <li>Ø Individuazione, classificazione delle violazioni secondo normativa vigente in materia</li> </ul>	<p>Garanzia dell'utenza e tutela dei minori</p>	<p>Generazione automatica di report secondo schemi Agcom (sia in formato PDF che Excel). Generazione automatica di Grafici</p>

Le metodologie di acquisizione ed analisi dei dati rilevati verranno messi a disposizione delle Università, firmatarie dell'Accordo di collaborazione a cui accede il presente Piano delle attività, al fine di condurre studi e ricerche, sperimentazioni di nuove tecnologie e metodiche atte a svolgere una completa e sicura elaborazione dei dati.

In particolare, si punta a sperimentare le grandi potenzialità delle intelligenze artificiali: Deep Learning, Machine Learning, Sentiment Analysis.

Inoltre, particolare attenzione sarà posta all'utilizzo di tecnologie del "Cloud Computing".

## **1.2 Implementazione di Progetti di Ricerca - Sperimentazione-Analisi su altre tipologie di Flussi Audio - Video Generate da altre fonti (RADIO-WEB-SOCIAL).**

Sulla scorta dell'elaborazione dei dati, che fungerà anche da supporto, ad opera dei ricercatori universitari per come previsto nell' Accordo di collaborazione, il Co.Re.Com. Calabria andrà a:

- sperimentare nuove e moderne forme di acquisizione, storage ed analisi dei flussi televisivi locali;
- intraprendere, contestualmente, una sperimentazione su flussi prodotti da altre fonti;
- elaborare il linguaggio naturale, l'analisi testuale e la linguistica computazionale per identificare ed estrarre informazioni soggettive da diverse fonti (Tv e radio locali, web, social).
- in particolare, a sperimentare le grandi potenzialità delle intelligenze artificiali: Deep Learning, Machine Learning, Sentiment Analysis, con particolare attenzione verso l'utilizzo di tecnologie di "Cloud Computing".

Nello stesso tempo, sempre nell'ottica di mettere in atto azioni sinergiche con le due Università, andrà a costituire per queste ultime un'opportunità, dal momento che la funzione di monitoraggio, svolta dal primo, unitamente ai dati rilevati nel tempo, costituirà un caso di studio, o meglio darà la possibilità di svolgere ricerca applicata. Inoltre, la messa a disposizione di specifiche competenze abbinate al proprio know-how tecnico da parte dei due Atenei, permetterà al Co.Re.Com. di accrescere l'esperienza fin qui acquisita e di rafforzare le proprie competenze tramite l'applicazione pratica di evolute tecniche di programmazione in grado di snellire, in maniera considerevole, le procedure di acquisizione, analisi ed elaborazione dei flussi e di interfacciarsi al meglio verso il Cloud.

Le predette azioni di intervento determineranno, quindi, progressi sia nell'ambito della ricerca ICT applicata all'acquisizione e all'analisi dei flussi audio video da fonti diverse, sia nell'ambito della semplificazione dei processi e la realizzazione di soluzioni open source.

Il Co.Re.Com. collaborerà in ogni fase del progetto costituendo un gruppo di lavoro che collaborerà con le Università secondo il presente Piano di attività condiviso, e verrà coadiuvato dal personale del Settore Informatico e Flussi Informativi del Consiglio regionale, compatibilmente con le attività di competenza del predetto Settore.

## **1.3 Aggiornamento – Formazione.**

La ricerca, la sperimentazione e le analisi condotte congiuntamente e in collaborazione con le due Università, saranno oggetto di un percorso di aggiornamento e di formazione continua che coinvolgerà il personale in servizio presso il Co.Re.Com. e il Settore Informatico e Flussi Informativi del Consiglio regionale della Calabria, anche mediante un processo di “Learning on the job”. L’obiettivo di questo percorso sarà quello di favorire sia l’acquisizione di competenze distintive, innovative e funzionali alle attività da svolgere, sia la sperimentazione applicata di metodologie di intelligenza artificiale per consentire approcci a nuove linee di ricerca in ambito di Cloud Computing.

Nell’ambito della cooperazione funzionale che si intende realizzare con l’Accordo a cui accede tale piano di attività, il Co.Re.Com. potrà collaborare con le due Università per la progettazione, l’organizzazione e la realizzazione di eventi informativi e/o seminari, manifestando la propria disponibilità a partecipare anche a pubblicazioni sugli esiti delle sperimentazioni e delle analisi condotte. Le tematiche innovative che verranno affrontate durante le attività di sperimentazione sono di seguito riportate:

- Il Cloud computing;
- Utilizzo dei servizi Cloud per l’acquisizione di flussi Audio Video;
- Utilizzo dei servizi Cloud per lo storage di flussi Audio Video;
- Utilizzo dei servizi Cloud per processing di flussi Audio Video e caratterizzazione;
- Approccio a Blocchi;
- Sicurezza del sistema informativo (GDPR e altro);
- Machine e Deep Learning;
- Big Data Analytics;
- Blockchain;

Le seguenti sezioni dettagliano gli avanzamenti tecnologici previsti durante la collaborazione tra il Co.Re.Com. e le Università di Reggio Calabria e Messina.

## **2. LINEE DI RICERCA DELLE UNIVERSITÀ.**

I gruppi di ricerca di Reggio Calabria e Messina possiedono una forte propensione alla ricerca di nuovi sistemi e nuove soluzioni informatiche. In questa collaborazione intravedono una concreta possibilità di testare concretamente quanto da loro studiato e prodotto. Dalle sezioni seguenti è possibile evincere le marcate competenze informatiche complementari dei due gruppi di ricerca coinvolti.



## 2.1 Università degli Studi di Messina (UniME).

### 2.1.1 Dipartimento di Scienze Matematiche e Informatiche, Scienze Fisiche e Scienze della Terra (MIFT).

Il gruppo di ricerca di Informatica “*Future Computing Research Laboratory*” (FCRLAB) del Dipartimento MIFT dell’Università degli Studi di Messina (UniME) ha marcata esperienza nell’ambito di ricerca dell’ICT.

Lo stesso da parecchio tempo affronta tematiche innovative, attraverso l’ideazione, la creazione e l’utilizzo di piattaforme ed architetture informatiche capaci di gestire grosse moli di dati (Big Data), estrapolando informazione a valore aggiunto (Big Data Analytics), attraverso lo sfruttamento opportunistico di risorse di calcolo di tipo cloud computing, edge computing e dell’Internet of Things IoT.

Il gruppo FCRLAB vanta una conoscenza dei sistemi informatici più innovativi.

Il coordinatore del gruppo è il Prof. Massimo Villari, che è co-autore di più di 170 pubblicazioni scientifiche in Cloud Computing (Cloud Federation), Distributed Systems, Wireless Network, Network Security, Cloud Security and Cloud and IoTs, di recente in Osmotic Computing.

Nell’anno 2014, il Prof. Villari è stato riconosciuto da uno studio indipendente pubblicato su IEEE Cloud Computing Transaction, Issue April 2014, come uno dei più attivi ricercatori scientifici al mondo (top 27 classification) in ambito di Cloud Computing.

Possiede un ottimo network internazionale con collaborazioni con parecchie università estere, (vedi TUW-Vienna AT, Cardiff Univ.- UK, New Castle-UK, Karlstadt-SE, Murcia-ES ); è editor di molte riviste internazionali in ambito ICT ed è spesso invitato come keynote speaker in molte conferenze internazionali. E’ uno dei fautori del nuovo paradigma dell’Osmotic Computing.

In diversi progetti europei in ambito ICT, è stato coordinatore di attività strategiche, come per il EU Project “RESERVOIR”, coordinando le attività di IT security.

Per i Progetti Europei “VISION-CLOUD” (F7P) e BEACON (H2020), ha rivestito il ruolo di architectural designer per UniME. E’ stato Scientific ICT Responsible nel progetto Europeo frontierCities, l’acceleratore dell’European Cloud-FIWARE su Smart Cities – Smart Mobility.

Attualmente è Responsabile Scientifico del progetto Europeo Cloud 4 Europe, PCP (Pre Commercial Procurement) gestito da AGID Italy.

## **2.2. Università degli Studi “Mediterranea” di Reggio Calabria.**

### **2.2.1 Dipartimento di Ingegneria Civile, Energia, Ambiente e Materiali (DICEAM).**

L'Elettrotecnica del DICEAM dell'Università degli Studi “*Mediterranea*” di Reggio Calabria vanta un'esperienza trentennale nel disegno, progettazione, implementazione e realizzazione software e hardware di sistemi di ML, specialmente basati sull'uso delle tecnologie neurali. Il coordinatore del gruppo è il Prof. Francesco Carlo Morabito, che ha presieduto la società italiana di reti neurali (SIREN) dal 2008 al 2014 ed è stato governatore di International Neural Network Society (INNS) dal 2000 al 2012.

Fra le numerose pubblicazioni internazionali del gruppo di ricerca facilmente reperibili sui database internazionali, quali Scopus e WoS. Si segnala, in particolare, il volume di prossima pubblicazione da parte di Elsevier/Academic Press, dal titolo “Artificial intelligence in the age of neural networks and brain computing”, di cui il Prof. Morabito è coeditore.

### 3. SPERIMENTAZIONE OPERATIVA.

In considerazione di una chiara esigenza di sperimentazione innovativa da parte del Co.Re.Com., riportata nel precedente paragrafo 1, le due Università di Reggio Calabria e Messina, individuano la possibilità di applicare, in un contesto reale, le conoscenze e le skills prodotte attraverso lo studio ed analisi di nuovi sistemi informatici, per lo più condotti in laboratorio o attraverso l'ausilio di simulatori.

In questo ambito, gli Enti di ricerca individuano una importante possibilità di applicare concretamente gli artefatti software, algoritmi e piattaforme ICT (Information and Communication Technology), sviluppati in situazioni circoscritte al fine di produrre nuovi risultati nel contesto della ricerca applicata.

In uno scenario internazionale, in cui l'analisi di casi di studio sta diventando sempre più decisiva per validare gli studi effettuati e dove, quindi, sta diventando sempre più importante l'applicabilità di soluzioni sperimentali in casi reali, l'interazione con il Co.Re.Com. per questi Enti diventa nevralgica.

D'altronde, l'Università italiana ha la necessità di attivare sempre più collaborazioni in tal senso. Basti pensare al programma di ricerca Europeo sotto il cappello H2020, che spinge i centri di ricerca internazionali ad applicare concretamente gli studi condotti dagli stessi centri di ricerca, al fine di validare quando sviluppato e fornire un tangibile riscontro di quanto prodotto nei laboratori ed in un contesto prettamente simulativo.

Le azioni RIA e IA di H2020 (Research and Innovation Action e Innovation Action) vanno fortemente in questa direzione e spingono i centri di ricerca europei ad uscire dai laboratori per fornire concreti risultati, misurabili dalle istituzioni e dai cittadini.

Spesso, nei progetti europei si parla di KPI (Key Performance Index) e Obiettivi Misurabili.

I primi sono rivolti ad una identificazione concreta di quali siano gli indici reali di innovazione (KPI) con le ricadute tecnologiche prodotte unitamente alle ricadute di business (spesso le soluzioni devono guardare alla creazioni di nuovi mercati es. *EU: Single Digital Market in IA initiative*). Gli altri obiettivi misurabili, spingono ad una reale presentazione dei risultati ottenuti caratterizzati da parametri tangibili e misurabili.

Non esiste attività di progettazione europea finanziata che non consideri questi concreti risvolti di applicabilità.

Il Co.Re.Com. rappresenta un laboratorio di studio, reale, molto particolare ed importante per i centri di ricerca coinvolti, visto che in virtù dei requisiti identificati, ha la necessità di gestire grosse mole di dati (ricadute di ricerca sperimentale in ambito di Big Data e MapReduce Algorithms) in maniera efficiente (ricadute di ricerca sperimentale in ambito di Cloud Computing), con la possibilità di automatizzare dei processi di analisi, catalogazione ed identificazioni di anomalie (ricadute di ricerca sperimentale in ambito di Machine Learning e Deep Learning).

## **3.1 Il CO.RE.COM. Calabria come test-case reale.**

Nell'ambito di questa attività di collaborazione congiunta, il Co.Re.Com. rappresenta un concreto validatore delle attività sperimentali dei gruppi di ricerca universitari coinvolti.

Infatti, questo organismo pubblico consente di poter implementare test e messa in funzione di un sistema informativo basato su tecnologie Cloud, in grado di immagazzinare e gestire grosse moli di dati (Big Data), utilizzando approcci basati su tag, offrendo servizi di accesso, elaborazione ed analisi di flussi video multimediali, in cui l'apprendimento automatico farà la differenza su quanto prodotto. Di seguito, vengono riportate tutte le tematiche innovative, che verranno affrontate durante l'attività di collaborazione congiunta, in cui si sperimenteranno e valideranno soluzioni, piattaforme e algoritmi nei vari ambiti di innovazione.

### **3.1.1 Il Cloud computing come base dei Sistemi Innovativi.**

Il Cloud computing rappresenta la naturale evoluzione dell'IT nel mondo enterprise, ponendosi come obiettivo l'astrazione dell'infrastruttura, vale a dire spostare l'attenzione dall'infrastruttura stessa alle informazioni (quello che realmente interessa l'utente finale), affermandosi come risposta alle esigenze sempre più "spinte" dei produttori e dei fruitori di dati. Le tecnologie Cloud sono usate per una ampia schiera di proposte, inclusa l'analisi dei dati, backup, archiviazione e per fornire software in modalità SaaS (Software as a service). In questo contesto, l'FCRLAB dell'Università di Messina ha elaborato e studiato soluzioni ad elevato potenziale innovativo nel settore dell'ICT basate su tecnologie Cloud.

Quando si parla di Cloud, generalmente lo si fa con riferimento all'utenza finale ad una serie di obiettivi, come, ad esempio, aumentare la flessibilità e liberare le risorse.

Con l'avvento delle applicazioni mobile e del Cloud computing, nei reparti IT si è imposta la necessità di lavorare con database flessibili e, soprattutto, scalabili. I database non relazionali o NoSQL hanno ricevuto un forte impulso di sviluppo da questa specifica esigenza, tanto da diventare quasi una scelta obbligata in determinati scenari d'uso, contrapponendosi ai database relazionali o SQL.

I database relazionali, come dice il termine stesso, si basano su relazioni univoche fra i dati. In pratica, tutti i dati da trattare sono memorizzati in strutture fisse, dette tabelle, dove sono posti l'uno di seguito all'altro a formare singole tuple o record (ossia le righe di ciascuna tabella).

Ogni dato contenuto in una tabella si trasforma in un'informazione vera e propria, soltanto se associato al suo corrispondente attributo, che rappresenta l'intestazione della colonna in cui il dato è memorizzato. Questa struttura obbliga la frammentazione delle informazioni tra differenti tabelle, anche quando i dati descrivono un medesimo oggetto. Per identificare le informazioni appartenenti a uno stesso oggetto, si utilizzano tutta una serie di operazioni logiche che come il JOIN,

basano i propri calcoli sulle chiavi esterne o foreign keys che mettono in collegamento una tabella all'altra.

Per questo motivo, utilizzare tabelle SQL troppo grandi è fortemente sconsigliato: sono di difficile gestione anche con sistemi computazionali molto potenti.

Questa peculiarità impone ai sistemi SQL di svolgere continue interrogazioni, in quanto i dati devono essere raccolti e combinati fra loro da differenti tabelle, anche per operazioni semplici, come l'inserimento, la cancellazione e l'aggiornamento. Questa forte frammentazione prescrive anche un rigido controllo sulle relazioni e sulla validità dei dati residenti nelle differenti tabelle in modo da preservare l'integrità del database, a scapito della flessibilità.

È anche per questo motivo che i database relazionali si basano su schemi tabellari (definiti schemi entità-relazione) predefiniti a monte e che difficilmente possono essere adattati a nuove situazioni, a meno di non correre il rischio di una corruzione dei dati.

I database non relazionali sono nettamente differenti da quelli SQL. Nei database *document-oriented*, i dati sono conservati in documenti, non in tabelle, per cui la prima differenza con i DB relazionali sta nel fatto che le informazioni non sono distribuite in differenti strutture logiche, ma vengono aggregate per oggetto in documenti la cui natura può essere di tipo Key-Value o Document Store basati su semantica JSON.

Ogni documento aggregato raccoglie tutti i dati associati a un'entità, in modo che qualsiasi applicazione possa trattare l'entità come oggetto e valutare in un sol colpo tutte le informazioni a essa correlate. In questo modo, si evitano anche i fardelli computazionali dovuti ai passaggi di aggregazione delle informazioni tipici del linguaggio SQL, in quanto tutti i dati necessari e corrispondenti a un medesimo oggetto sono già disponibili in un unico documento. L'assenza di tabelle permette ai database non relazionali di essere *schemeless*, ossia privi di un qualsiasi schema definito a priori e questa caratteristica conferisce ai DB NoSQL un altro vantaggio non trascurabile. In un sistema aziendale consolidato, non è possibile operare un passaggio completo da un sistema di gestione dati ad un altro per due ragioni: 1) le tecnologie SQL e No-SQL presentano entrambe vantaggi di cui si vuole usufruire, 2) il passaggio a tecnologie innovative deve essere graduale per non condizionare l'efficienza del sistema informativo esistente. Per i suddetti motivi, è possibile adottare un innovativo sistema di gestione dati ibrido, che gestisca con tecnologie diverse esigenze differenti.

A seguire vengono riportati i possibili campi di sperimentazione in cui si ritiene che i centri di ricerca possano validare le proprie attività di ricerca associate alle specifiche esigenze del Co.Re.Com..

### **3.1.2 Sperimentazioni legate all'utilizzo dei servizi Cloud per Audio - Video Acquisition.**

Il Co.Re.Com. offre degli ottimi spunti nella gestione e presenta interessanti sfide nell'ambito del Audio Video acquisition.

Per Audio Video acquisition si intende il processo di ricezione e transcodifica dei flussi audio-video.

Esso ha la necessità di realizzare un sistema di acquisizione flessibile capace di scegliere qualunque sorgente di acquisizione (SAT, Terrestre, IP, ecc.) e di interpretarla e transcodificarla con le più moderne tecnologie di codec-video-compression.

Quindi, input selezionabile, transcodifica ed output ad hoc, pronto all'uso dall'applicazione a valle del sistema di Video acquisition.

Allo stato attuale l'acquisizione base prevede, in prima battuta, un tipo di ricezione da Tuner TV con Antenna di tipo Digitale Terrestre (DTTv1, successivamente DTT v2). In fasi successive si prevede l'analisi di segnali radiofonici e Satellitari (DVBSAT1 (S1) e DVBSAT2 (S2)). Infine, le sorgenti possibili saranno di tipo Web, di due tipologie: a) da sito web emittente locale e b) da web youtube/facebook o altra piattaforma utilizzata.

### **3.1.3 Sperimentazioni legate all'utilizzo dei servizi Cloud per storage dei flussi audio-video.**

Ad oggi, al Co.Re.Com., lo storage dei file avviene attraverso l'ausilio di sistema di conservazione di tipo NAS (Network Attached Storage), che rappresentano il sistema embrionale di base per sistemi di Storage presenti in reti LAN.

Si potrebbero attuare nuovi meccanismi di conservazione connotati da principi di affidabilità e backup dei dati. Il cloud storage, per mezzo della sua unità di conservazione definita Object Storage (ad esempio Amazon S3), risponde pienamente a questi requisiti.

Da qui, si deve prevedere la possibilità di testare la migrabilità del dato all'interno di cloud storage private, pubbliche o ibride, senza effort di migrazione particolari, se non un semplice copia ed incolla.

Nella conservazione del flusso video, a valle della transcodifica, è necessario arricchire lo stesso con dati embedded in sovraimpressione utili a definire la data ed il tempo (time: orologio atomico) di acquisizione unitamente alla sorgente di acquisizione. Vista l'importanza della gestione temporale (e sovraimpressione dello stesso dato), si ha la necessità prioritaria di gestire l'esattezza del dato time e, da qui, scaturisce la necessità di utilizzare più sorgenti tempo diverse (Radio Signal es. da Francoforte - DCF77 - e dato IP da NTP servers.). La sincronizzazione oraria tra diverse fonti è un'operazione complessa che richiede studi specifici. I dipendenti del Co.Re.Com., coadiuvati dal gruppo di ricerca FCRLAB, effettueranno ricerca applicata al fine di implementare tecniche innovative di sincronizzazione oraria.

### **3.1.4 Sperimentazioni legate all'utilizzo dei servizi Cloud per processing dei flussi audio-video e caratterizzazione.**

Il Co.Re.Com. ha la necessità di affrontare delle questioni significative nell'ambito della caratterizzazione dei flussi audio-video da parte di AgCom.

In tale scenario, i flussi, una volta acquisiti e conservati in forma nativa, devono essere elaborati e processati per l'arricchimento delle acquisizioni fatte nel rispetto della normativa vigente e delle direttive dell'Autorità.

Il Cloud processing distribuito, permette di segmentare i flussi di ingresso e dividerli per un post-processing molto più efficiente. In questa fase si prevede un arricchimento guidato dall'utente, che, aiutato da una interfaccia, dà la possibilità via web-cloud di aggiungere tutti i metadati di arricchimento, sia per un ulteriore post-processing (Big Data Analytics per statistiche di interventi/pubblicità/ecc determinatosi durante le mandate in onda) e sia per la produzione documentale (report pdf) necessaria a supportare eventuali ricorsi e/o controversie. Ad oggi, non esiste alcun prodotto commerciale che, utilizzando il cloud, possa soddisfare queste esigenze, perciò, il Co.Re.Com. ed il gruppo di ricerca FCRLAB effettueranno ricerca congiunta al fine di progettare ed implementare un prototipo software che possa includere queste funzionalità.

### **3.1.5 Applicare un nuovo Approccio a Blocchi.**

Il sistema che si immagina di sperimentare, permetterà di progettare e realizzare servizi utilizzando un approccio grafico con struttura a blocchi (vedi Image 1), così da supportare un qualunque operatore non esperto per la creazione di nuovi servizi.

L'interfaccia visuale, ad esempio, permetterà di aggiungere più sorgenti di acquisizione. Gran parte dei blocchi saranno configurati con meccanismi di tipo API Web-Cloud di denominazione RESTful. L'intero sistema funzionerà tenendo conto di innumerevoli interfacciamenti con sistemi esistenti (come gmail, twitter, telegram, nextcloud, ecc.). Il Co.Re.Com. ed il gruppo di ricerca FCRLAB sperimenteranno ed effettueranno ricerca congiunta su tecnologie opensource capaci di ottemperare all'esigenza di un'interazione semplificata da parte degli utenti ed operatori.

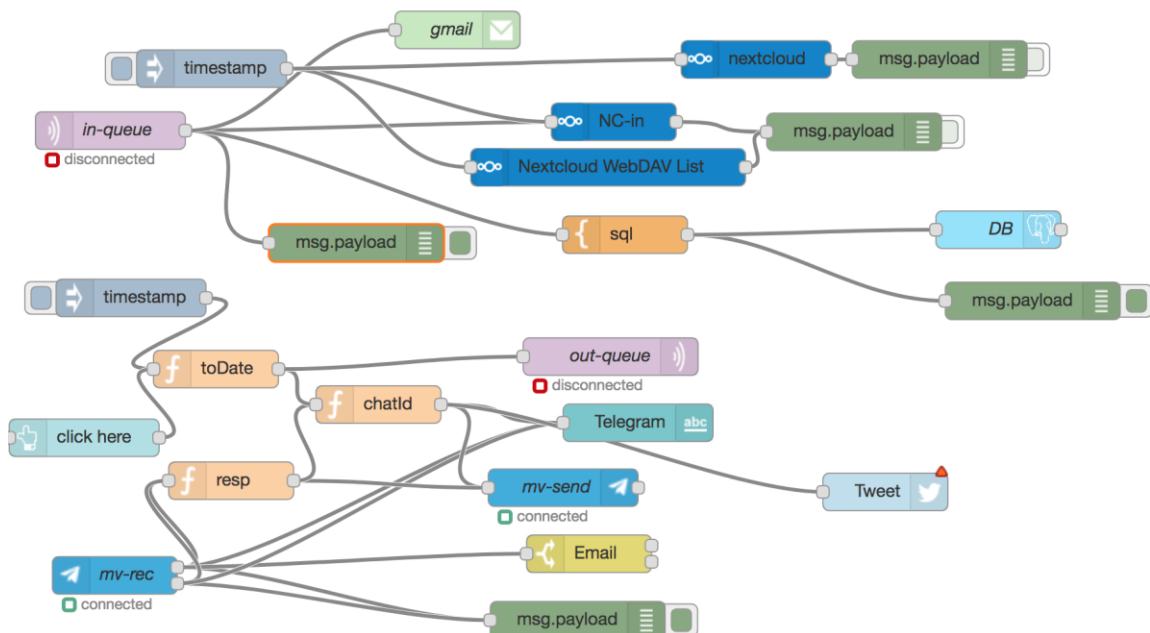


Image 1: Sistema a Blocchi Visuali

### 3.1.6 Sicurezza del sistema informativo.

Lo scorso 25 Maggio 2018 la Commissione Europea ha reso obbligatorie delle regolamentazioni nell'ambito della gestione del Data Privacy degli utenti che interagiscono con sistemi informativi digitali. Per preservare la Privacy degli utenti è necessario mettere in atto delle procedure capaci di identificare gli utenti (Identity Management), controllare le attività che svolgono dopo aver acceduto al sistema (LOGs di sistema), provvedendo alla corretta gestione degli accessi (Access Management).

Il Co.Re.Com. ed il gruppo di ricerca FCRLAB effettueranno ricerca avanzata al fine di progettare e sviluppare un sistema con connotazioni di base di tipo GDPR (native GDPR compliant). Si prevede lo studio, la progettazione e la creazione di un IAM (Identity and Access Management) che attraverso dei sistemi a Token (OAuth2.0) permetterà di rendere sicure tutte le API Web-Cloud di denominazione RESTful, anche attraverso l'ausilio di certificati x509v3 utili ad attivare canali SSL/TLS based. L'Access Management verrà garantito utilizzando le più moderne tecnologie di Attribute Based Encryption.

### 3.1.7 Applicazione del Machine e Deep Learning.

Il Machine Learning (ML, Apprendimento Automatico), si riferisce ad un insieme di algoritmi implementati attraverso opportuni codici su specifici framework, attraverso cui è possibile sviluppare sistemi autonomi con capacità adattiva e "intelligenti", in quanto in grado di imparare sulla base delle informazioni a disposizione e,



conseguentemente, prendere decisioni o effettuare previsioni con un elevato tasso di accuratezza.

Esistono diversi paradigmi di apprendimento automatico: supervisionato, non supervisionato, semi-supervisionato, per rinforzo.

Nell'apprendimento supervisionato, il computer riceve in ingresso sia il dato originale che il risultato desiderato ("target"). Il sistema ha l'obiettivo di trovare una relazione generale ingresso-uscita tale che, in presenza di un nuovo dato in ingresso si possa ottenere il risultato corretto, cioè di generalizzare; al contrario, nell'apprendimento non supervisionato, i dati di ingresso non sono etichettati (la classe di appartenenza non è nota a priori). In tal caso, il sistema di apprendimento automatico ha lo scopo di trovare una relazione statistica (se esiste), ovvero un insieme di correlazioni, all'interno dei dati di ingresso forniti; nell'apprendimento semi-supervisionato, solo una minima parte dei dati di ingresso è etichettata; nell'apprendimento per rinforzo, il sistema interagisce con l'ambiente in cui è immerso ed impara a modificare il proprio comportamento in risposta a una qualche forma di "ricompensa" ricevuta. Non esiste un algoritmo migliore di un altro, la scelta della metodologia di apprendimento più idonea, dipende dallo specifico problema da affrontare.

Le tecniche di machine learning sono oggi ubiquo, ampiamente utilizzate in diverse applicazioni industriali e manageriali. In particolare, le tecniche di apprendimento approfondito, note come deep learning (DL), sono la base dei più recenti sistemi IT e, più in generale, della cosiddetta Industria 4.0.

Le tecniche Deep Learning, utilizzano grandi set di dati (Big Data) etichettati e sono principalmente basate sulla teoria delle reti neurali artificiali. Esempi di reti profonde sono: le convolutional neural network, stacked autoencoder, deep belief network: esse permettono di estrarre/apprendere le caratteristiche (features) più significative automaticamente dai dati di ingresso, in modalità cioè data-driven, evitando il calcolo manuale, tipico di un approccio standard.

In questo contesto, le tecniche di deep learning trovano applicazione soprattutto nell'analisi ed elaborazioni dei dati immagazzinati su piattaforme Cloud, in quanto apre l'accesso ad una quantità di informazioni ancora più grande.

L'integrazione di algoritmi "deep" ad infrastrutture Cloud sta crescendo negli ultimi anni e risulta essere la strategia ottimale: permette, infatti, di sviluppare sistemi più accurati e di estrarre le caratteristiche più importanti da un flusso di informazioni elevato.

Le attività, oggetto del presente accordo di collaborazione, riguarderanno la sperimentazione di un sistema basato sia su tecnologie Cloud, che deep learning. In particolare, l'attività comprenderà la progettazione, lo sviluppo ed il test di reti neurali profonde a supporto, per il processamento dei flussi video e l'estrazione delle feature più significative per garantire una più accurata (e rapida) analisi dei dati in piattaforme Cloud.

Le suddette attività verranno svolte congiuntamente da UniRC, UniMe e Co.Re.Com..

### 3.1.8 Interagire con grosse moli di dati: Big Data.

Fra le tecnologie abilitanti individuate dal piano Industria 4.0, sono incluse le soluzioni per il manifatturiero avanzato (per esempio, per la robotica collaborativa e interconnessa) e i “Big Data and Analytics” (per esempio, per l'estrazione di informazioni d'interesse da ampie basi di dati per l'ottimizzazione dei processi produttivi).

Entrambe queste tecnologie sono largamente supportate dalle tecniche di apprendimento automatico (Machine Learning, ML), anche di tipo approfondito, in caso di Big Data (Deep Learning, DL).

Queste nuove tecnologie, ormai utilizzate dai maggiori gruppi industriali manifatturieri, oltre che dai colossi dell'informatica e delle comunicazioni, consentono di estrarre dai dati disponibili (quindi, in modalità data-driven) i modelli sottostanti da cui vengono generati i dati acquisiti da sistemi di monitoraggio e controllo.

Da un punto di vista strettamente matematico, si tratta di stimare la funzione densità di probabilità di un modello multi-input e multi-output (MIMO) a partire da un insieme di dati sperimentali affetti da rumore. Il problema è intensivo dal punto di vista computazionale, oltre che mal posto, poiché i dati sono spesso insufficienti per garantire una stima a bassa varianza dei parametri liberi del sistema.

Inoltre, in apparecchiature industriali complesse, i dati acquisiti vengono aggiornati continuamente in real-time, modificando quindi il modello sottostante in continuità.

Le tecniche di ML e DL, basate sulle tecniche di apprendimento automatico, consentono di affrontare con successo le problematiche sopra esposte.

Alcuni dei benefici attesi dall'uso di tali tecnologie sono, per esempio, la riduzione degli errori di produzione, il monitoraggio anche in tempo reale della produzione, qui intesa in termini di flussi video da analizzare, la riduzione dei consumi energetici anche mediante l'utilizzo di sensori multipli e disomogenei connessi in rete.

I vantaggi rispetto alla modellistica tradizionale possono essere riassunti nei punti seguenti:

1. Costruzione del modello a partire dai dati disponibili;
2. Aggiornamento real-time del modello;
3. Possibilità di implementazione di metodologie diverse (ottimizzazione evolutiva, fuzzy logic, reti neurali, modelli di ant colony, etc) in un unico modello;
4. Riduzione automatica della rumorosità dei dati acquisiti (es. jitter);
5. Costruzione di modelli MIMO a partire dalle variabili definite dall'utente;
6. Estrazione di relazioni multidimensionali fra le variabili d'interesse per l'utente.

Il Co.Re.Com. consente di attuare nuove metodologie di analisi dei dati.

Poiché il monitoraggio dei flussi televisivi o di altra origine delle attività fa riferimento a un flusso continuo di dati, dal complesso disponibile di tali dati (big data) è potenzialmente possibile estrarre delle correlazioni statistiche tra input (flussi video)

e output (definiti da normative e direttive AgCom), che consentono di identificare le sotto-relazioni che sono significative per la caratterizzazione e l'analisi dei flussi.

Tali relazioni sono in alcuni casi non lineari, il che giustifica l'utilizzo di dispositivi quali le reti neurali per la determinazione di tali correlazioni.

L'obiettivo comune a Co.Re.Com., UniRC è quello di effettuare ricerca congiunta ed applicata al fine di determinare, a partire dalla disponibilità dei dati in possesso del Co.Re.Com. (resi disponibili per la ricerca nelle modalità ritenute opportune), dei modelli di identificazione, caratterizzazione, automazione e controllo dei flussi video secondo le direttive AgCom.

Fra le tecniche proposte, il DL di tipo convolutivo e con tecniche di compressione sparsa appaiono maggiormente significative e innovative.

### **3.1.9 Guardare dentro ai dati: Big Data Analytics.**

Il modulo di Big Data Analytics avanzato, basato sul Machine Learning, riguarda l'intero processo di supporto alla classificazione di sequenze video (es. identificazione dei break, delle pubblicità, rispetto par condicio, ecc.). Il modulo rappresenta un supporto avanzato agli operatori, in cui degli automi intelligenti, progettati sulla base dei più avanzati modelli di Machine Learning, permettono delle funzionalità di riconoscimento della scena all'interno dei video sotto processing in maniera completamente automatizzata ed innovativa.

Questi sistemi, verranno implementati utilizzando tecniche di deep learning introdotte precedentemente, al fine di ottenere le massime prestazioni. E' previsto quindi l'arricchimento automatico delle acquisizioni fatte anche in ottica delle normative vigenti in materia sia delle direttive AgCom.

L'analisi di flussi video, richiede un elevato livello di flessibilità a causa di una rete continuamente aggiornata e modificabile. La continua crescita di nuove tipologie di video streaming abbinata alla emergenza di standard di compressione ad alta qualità e alla variegata tipologia di contenuti video, impone la ricerca di metodi generali e scalabili.

Per di più, nel caso specifico, l'analisi e caratterizzazione dei flussi video deve soddisfare le specifiche AgCom. Per tali ragioni, è necessario progettare dei sistemi di Quality Assessment (QA) scalabili e adattivi. Inoltre, problematiche statistiche quali la perdita di pacchetti, il jitter e i ritardi rendono talora imprevedibile, se non con tecniche statistiche, l'analisi stessa dei flussi.

I nuovi contenuti e framework video pongono sempre più stringenti vincoli sui metodi di QA implementati. QA deve essere accurata, in real-time, adattabile a nuovi contenuti e scalabile con le dimensioni dei flussi disponibili. Si propone una metodologia innovativa di QA automatizzata e efficiente dal punto di vista computazionale. In principio, essa consente una analisi accurata, real-time (online) di QA che sia adattiva e scalabile. Tecniche offline di "deep unsupervised learning" verranno utilizzate lato server ed eventualmente lato client. La presenza di un

modulo di natura unsupervised learning assicura un metodo adattivo e scalabile con il numero di video.

In particolare, QA in streaming video, richiede una metrica di qualità che sia adeguata per il real-time e per una vasta gamma di tipi di video e condizioni di rete.

Le prestazioni del sistema nel suo complesso vengono misurate attraverso opportune metriche costruite attraverso “features” temporali e spaziali estratte dai video analizzati. Inoltre, l’analisi delle variazioni temporali della sequenza video può essere analizzata in opportuni domini trasformati (wavelet, ...). Nel corso del progredire, si farà uso di tecniche di apprendimento automatico e di analisi multivariata appropriate per i big data. La parte di training del processo viene sviluppata offline attraverso approcci “cognitivi”.

A causa della necessità di far fronte alle esigenze di scalabilità e adattività, Co.Re.Com, UniRC ed UniME effettueranno ricerca avanzata per utilizzare tecniche di apprendimento unsupervised e semi-supervised, ad esempio le reti convolutive, complessivamente indicate come Deep Learning.

### **3.1.10 Nuove soluzioni attraverso l’uso della Blockchain.**

La Blockchain (traducibile in italiano con “catena di blocchi”), è una tecnologia che permette la creazione e gestione di un grande database distribuito per la gestione di transazioni condivisibili tra più nodi (server di ciascun partecipante) di una rete.

Ha una storia pluridecennale, ma è solo dal 2009 che ha trovato una sua definitiva collocazione nel mondo dell’hi-tech; merito di un’intuizione di Satoshi Nakamoto, che ha deciso di fondare il funzionamento dei Bitcoin su di essa.

Si tratta di un database strutturato in Blocchi (contenenti più transazioni) che sono tra loro collegati in rete in modo che ogni transazione avviata sulla rete debba essere validata dalla rete stessa nell’analisi di ciascun singolo blocco.

La Blockchain è, in sostanza, un registro pubblico nel quale vengono archiviati in ordine temporale e in modo sicuro, verificabile e permanente, le transazioni che avvengono tra due utenti appartenenti a una stessa rete.

I dati relativi agli scambi sono salvati all’interno di blocchi crittografici, collegati in maniera gerarchica l’uno all’altro. Si viene così a creare un’infinita catena di blocchi di dati che consente di risalire e verificare tutte le transazioni fatte.

La funzione primaria di una blockchain è, dunque, di certificare transazioni tra persone. Nel caso dei Bitcoin la blockchain serve a verificare lo scambio di criptovaluta tra due utenti, ma si tratta solo di uno dei tanti possibili utilizzi di questa struttura tecnologica. In altri settori, la blockchain può certificare lo scambio di titolo e azioni; operare come fosse un notaio e “vidimare” un contratto o rendere sicuri e non alterabili i voti espressi tramite votazione online.

Uno dei maggiori pregi della blockchain è l’elevato grado di sicurezza che garantisce.

Di fatto, una volta che una transazione viene certificata e salvata all'interno di uno dei blocchi della catena non può più essere modificata, né manomessa.

Questa tecnologia è importante in termini di:

- **affidabilità** perché consente a tutti i partecipanti diretti di avere una parte di controllo dell'intera catena e quindi non vi è un unico server che controlla l'intero processo;
- **trasparenza** infatti le transazioni effettuate attraverso la blockchain sono visibili a tutti i partecipanti, garantendo così trasparenza nelle operazioni;
- **convenienza** in quanto vengono meno interlocutori di parti terze, necessari in tutte le transazioni convenzionali che avvengono tra due o più parti (ovvero le banche ed altri enti simili);
- **solidità** perché le informazioni contenute nella blockchain sono tutte più solide ed attendibili, per il fatto che non si possono alterare e quindi restano così come sono state inserite la prima volta;
- **irrevocabilità**: con la blockchain è possibile effettuare transazioni irrevocabili, e allo stesso tempo più facilmente tracciabili. In questo modo si garantisce che le transazioni siano definitive, senza alcuna possibilità di essere modificate o annullate;
- **digitalità** in quanto con la blockchain tutto diventa virtuale. Grazie alla digitalizzazione, gli ambiti applicativi di questa nuova tecnologia diventano molteplici come ad esempio: finanza e banche, assicurazioni, Pagamenti digitali, Agrifood, Industry 4.0 IoT, Sanità, pubblica Amministrazione e Retail.

Nell'ambito della proposta progettuale la Blockchain potrebbe essere adottata tutte le volte che è necessario certificare delle azioni compiute all'interno della piattaforma stessa. Questo sistema è compatibile le normative vigenti in materia di monitoraggio televisivo e con le direttive AgCom.

Con riferimento alla sezione Blockchain, l'impiego di tecniche di ML e DL potrà essere altresì utile. Infatti, anche un sistema basato sulla tecnologia blockchain potrebbe essere soggetto ad accessi non autorizzati da parte di enti esterni, mettendo a rischio l'integrità e la riservatezza delle informazioni trasmesse.

A questo proposito, Co.Re.Com. UniRC ed UniMe dovranno, attraverso attività di ricerca congiunta, studiare, progettare ed implementare sistemi "deep learning oriented" per il riconoscimento e rilevamento di possibili attacchi hacker, garantendo quindi una maggiore sicurezza informatica.

## **4. RISORSE UMANE IMPIEGATE NEL PROGETTO.**

### **4.1 Risorse umane impiegate dalle Università.**

Per l'attività sperimentale e di ricerca applicata descritta nei paragrafi precedenti, le due Università intendono darvi corso con cinque borse di studio per attività di ricerca *post lauream* per laureati con profilo di sviluppatore (3 unità per l'UniMe e 2 unità per l'UniMed).

L'attività di coordinamento, che si protrarrà per tutta la durata del progetto per una costante supervisione degli obiettivi da raggiungere durante le varie fasi di progettazione, sviluppo ed implementazione, e il ruolo di responsabili scientifici del progetto, saranno svolti, per l'UniMe, dal Prof. Massimo Villari, per l'UniMed dal Prof. Francesco Carlo Morabito.

Il Co.Re.Com. supporterà parte dei costi finanziando le borse di studio, nonché le spese coerenti annesse (es. acquisto hardware, software, ecc.) che sosterranno i due Atenei, per il tramite dei rispettivi Dipartimenti, in vista della gestione delle prime. Il Co.Re.Com. rimane estraneo da qualsiasi rapporto economico con i borsisti, per cui gli unici referenti di quest'ultimi saranno i responsabili scientifici delle due Università.

## **4.2 Risorse umane impiegate dal Co.Re.Com. Calabria.**

Per l'attività sperimentale e di ricerca applicata descritta nei paragrafi precedenti, si impiegherà il personale in servizio presso i competenti Uffici del **Co.Re.Com.**, che verrà coadiuvato da quello incardinato al Settore Informatico e Flussi Informativi del Consiglio regionale, compatibilmente con le attività di competenza del predetto Settore.

L'attività di coordinamento, che si protrarrà per tutta la durata del progetto per una costante supervisione degli obiettivi da raggiungere durante le varie fasi di ricerca, progettazione, sviluppo ed implementazione, e il ruolo di responsabile del progetto, sarà svolto, per il Co.Re.Com. Calabria, dal dirigente del Settore o un suo delegato.

## **4.3 Interazione gruppi di lavoro.**

1. Le Parti concorderanno le opportune modalità di informazione e di comunicazione utili ad agevolare l'accesso dei componenti del Gruppo di lavoro alle fonti di informazione istituzionale, rientranti nelle disponibilità di ciascuna di esse. Al fine di una migliore interazione tra i tre gruppi di ricerca (Co.Re.Com., UniMed, UniMe), verranno allestite nei locali del Co.Re.Com. specifiche postazioni atte ad ospitare i borsisti, i quali vi si recheranno con una frequenza di una volta alla settimana. Tale frequenza potrà variare in base alle esigenze del progetto, previo accordo tra i responsabili scientifici. Parimenti, le due Università predisporranno, presso le proprie sedi, delle postazioni che potranno essere utilizzate dal personale del Co.Re.Com., in base alle esigenze del progetto.

2. Le Parti predisporranno dei report semestrali sulle azioni condotte, di cui l'ultimo dovrà contenere le risultanze finali dell'attività della collaborazione intercorsa tra di esse.

#### 4.4 Gantt delle attività.

Nella tabella successiva, è presente il diagramma di Gantt.

Per ogni attività sono stati rappresentati il mese di inizio, il mese di fine ed i Mesi Uomo necessari a svolgere l'attività stessa, tenendo conto, tuttavia, di un periodo minimo di durata pari a 12 mesi.

N	Descrizione Attività	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	Attività (Mesi Uomo)
1	Coordinamento													4
2	Analisi dei Requisiti													2
3	Design e Setup Infrastruttura Cloud per Computation, Storage e Networking													5
4	Design e Setup di servizi orientati all'acquisizione e trasferimento dei flussi video													5

<b>N</b>	<b>Descrizione Attività</b>	<b>M1</b>	<b>M2</b>	<b>M3</b>	<b>M4</b>	<b>M5</b>	<b>M6</b>	<b>M7</b>	<b>M8</b>	<b>M9</b>	<b>M10</b>	<b>M11</b>	<b>M12</b>	<b>Attività (Mesi Uomo)</b>
<b>5</b>	<b>Design, Sviluppo e Setup di moduli Back-End per la gestione di Flussi Video e Processing Big Data</b>													<b>11</b>
<b>6</b>	<b>Design, Sviluppo e Setup di moduli Front-End Web per gestione Flussi Video e configurazion e Workflow.</b>													<b>7</b>
<b>7</b>	<b>Design, Sviluppo e Setup di moduli Front-End per la creazione di reportistica complessa.</b>													<b>6</b>
<b>8</b>	<b>Design e setup di algoritmi e architetture di ML/DL per l'elaborazione dei flussi video</b>													<b>7</b>



<b>N</b>	<b>Descrizione Attività</b>	<b>M1</b>	<b>M2</b>	<b>M3</b>	<b>M4</b>	<b>M5</b>	<b>M6</b>	<b>M7</b>	<b>M8</b>	<b>M9</b>	<b>M10</b>	<b>M11</b>	<b>M12</b>	<b>Attività (Mesi Uomo)</b>
<b>9</b>	<b>Design e setup di algoritmi e architetture di ML/DL per la classificazione e di sequenze video</b>													<b>7</b>
<b>10</b>	<b>Design, sviluppo e setup di moduli specifici di ML/DL per l'estrazione automatica delle informazioni d'interesse da flussi</b>													<b>7</b>
<b>11</b>	<b>Design e setup di algoritmi e architetture di ML/DL per il riconoscimento/rilevamento di possibili attacchi hacker</b>													<b>7</b>

## **5. Dotazione Hardware- progetto sperimentazione.**

### **5.1 Ricevitore segnale.**

Da un'indagine di mercato si è visto che il Terracue DXP-380D è uno dei dispositivi più avanzati sul mercato, esso è dotato di otto sintonizzatori indipendenti, che possono essere ingressi DVB-S / S2, T2 / T / C.

### **5.2 Unità di storage locale: mini cloud privato.**

Per quanto riguarda lo spazio di archiviazione, si è pensato di utilizzare il cloud privato. In particolare si è considerato uno storage ridondato di 40TB capace di contenere le registrazioni h24 di 8 canali in alta risoluzione per un arco temporale di oltre 90 giorni.

- a. Storage locale: 2 NAS Synology RackStation RS818RP+ ognuno equipaggiato con
  - 1. 4x HDD 10TB ST10000NM0016 - 1TT101 (8 in totale)
  - 2. 2x Modulo RAM RAM1600DDR3L-8GBx2 (16GB) (2 in totale)
  - 3. Guide per rack: 2x RKS1317
- b. Computazione locale: 1 Storage server Dell PowerEdge R630 equipaggiato con:
  - 1. Processori: 2x Intel® Xeon® E5-2630 v4 2.2GHz,25M Cache,8.0 GT/s QPI,Turbo,HT,10C/20T (85W) Max Mem 2133MHz;
  - 2. RAM: 2x 16GB RDIMM, 2666MT/s, Dual Rank;
  - 3. Sistema operativo: Citrix XenServer 7.0 Factory Installed on HDD, Unlicensed;
  - 4. Configurazione RAID: C3 - RAID 1 for H330/H730/H730P (2 HDDs or SSDs);
  - 5. Disco Rigido:2x 960GB Solid State Drive SATA Read Intensive TLC 6Gbps 512e 2.5in Hot-plug Drive, S4500;
  - 6. Raffreddamento: Fresh Air Cooling
  - 7. Alimentatore: Dual, Hot-plug, Redundant Power Supply (1+1), 1100W;
  - 8. Guide per rack: ReadyRails™ Sliding Rails With Cable Management Arm
- c. Mikrotik rb3011

## 5.3 Unità di computazione remota: cloud pubblico Amazon

Per le operazioni di computazione e trascodifica si sono considerate 2\* macchine virtuali su Amazon "t3.large" aventi cioè le seguenti caratteristiche:

CPU: 2

RAM: 8 GB

HDD: 60 GB

## 5.4 Unità di calcolo locale per Machine Learning

Per le elaborazioni è necessario un server ad alte prestazioni e un sistema di salvataggio temporaneo dei modelli interconnessi da switch di rete.

A) Test modelli :1 PowerEdge R730 equipaggiato con:

1. Processore: 1x Intel® Xeon® E5-2660 v3 2.6GHz,25M Cache,9.60GT/s QPI,Turbo,HT,10C/20T (105W) Max Mem 2133MHz;
2. RAM: 16GB RDIMM, 2666MT/s, Dual Rank;
3. Sistema operativo: None;
4. Configurazione RAID: C1 - No RAID for H330/H730/H730P (1-16 HDDs or SSDs);
5. Dischi Rigidi:4 x 300GB 15K RPM SAS 12Gbps 2.5in Hot-plug Hard Drive;
6. Raffreddamento: Deepcool Captain
7. Alimentatore: Dual, Hot-plug, Redundant Power Supply (1+1), 1100W a efficienza Platinum;
8. Guide per rack: ReadyRails™ Sliding Rails With Cable Management Arm;
9. Broadcom 5720 QP 1Gb Network Daughter Card;
10. PERC H730 RAID Controller, 1GB NV Cache;
11. Slot configuration #2: 2 x Full Length, Full Height, PCIe Gen3 x16 (x16 connector)
12. Chassis with up to 8, 2.5" Hard Drives
13. 2 x Dell GPU Installation Kit;
14. 2 x GPU Nvidia Tesla V100 PCIe Full Height
15. 2 x Kit di installazione GPU Nvidia Tesla V100
16. Cavi, raffreddamento liquido per cpu e minuteria