

I-Media-Cities, una piattaforma multidisciplinare per l'analisi e l'annotazione di materiale video

Simona Caraceni, Michele Carpenè, Mattia D'Antonio, Giuseppe Fiameni, Antonella Guidazzoli, Silvano Imboden, Maria Chiara Liguori, Margherita Montanari, Giuseppe Trotta, Gabriella Scipione

Cineca

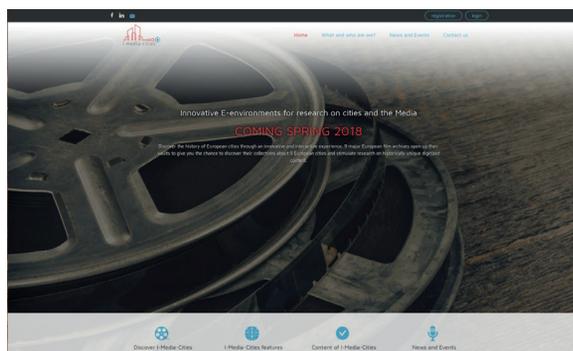
Abstract. Il progetto europeo I-Media-Cities ha come obiettivo quello di dare nuove possibilità di accesso al Patrimonio audiovisivo delle cineteche e archivi europei. Grazie ad innovativi strumenti di ricerca, decodifica automatica di immagini in movimento ed annotazioni automatiche di metadati, diventa possibile esplorare le città da prospettive finora inedite.

Keywords. Accesso libero ai dati, Condivisione dei dati e cloud storage, Research Support Issues, Conservazione dei dati.

Introduzione

I-Media-Cities è un progetto finanziato dalla Comunità Europea che coinvolge 9 archivi cinematografici europei, 5 istituti di ricerca, 2 fornitori tecnologici e uno specialista di modelli di business digitali (<https://imediacities.eu/>). L'obiettivo principale è la messa in condivisione, l'accesso e la valorizzazione dei contenuti audiovisivi presenti negli archivi per favorire la ricerca nel settore delle Digital Humanities. Il progetto ruota attorno alla storia delle città coinvolte e prevede la realizzazione di una piattaforma innovativa attraverso la quale è possibile raccogliere, analizzare, integrare e condividere l'enorme quantità di opere audiovisive provenienti dagli archivi - dalla fine del XIX secolo in poi - che descrivono le città in tutti gli aspetti, tra cui la trasformazione fisica e le dinamiche sociali. La piattaforma prevede la realizzazione di diversi e-environment che verranno utilizzati da ricercatori ed innovatori per la ricerca e per altri scopi creativi

Fig. 1
Home page del sito I-Media-Cities



(figura 1). Ciò consentirà nuovi approcci alla ricerca nell'ambito di Digital Humanities, Scienze Sociali ed industrie creative.

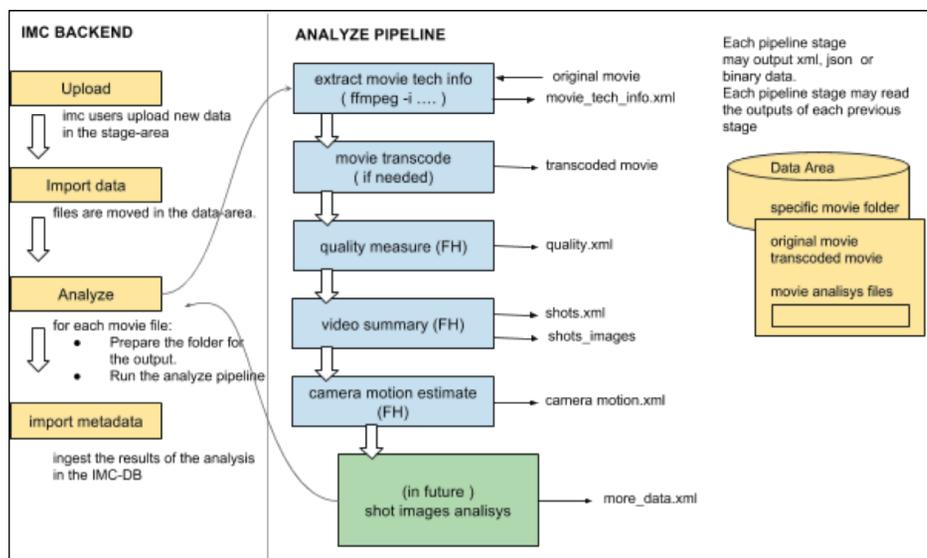
Il progetto triennale - attualmente alla fine del suo secondo anno - ha raggiunto una importante milestone nello sviluppo della piattaforma che rappresenta la spina dorsale di tutto il progetto. L'aspetto innovativo consiste nella possibilità di eseguire annotazioni automatiche sui contenuti multimediali che sono già stati arricchiti, a livello di documento, con metadati provenienti dagli archivi e che verranno ulteriormente annotati manualmente.

La piattaforma riceve i contenuti in forma di video e immagini, a cui vengono associati i metadati forniti in file XML dagli archivi audiovisivi. Una volta che i contenuti e i metadati vengono caricati all'interno della piattaforma, diversi strumenti di analisi video automatici, orchestrati attraverso una pipeline, li analizzano ed estraggono informazioni e metadati, in formato XML o in JSON, a livello di fotogramma o a livello di scena.

1. Annotazione e tagging automatico

La piattaforma IMC deve essere in grado di estrarre le annotazioni automatiche tramite strumenti di analisi video forniti dal partner tecnico Fraunhofer per Digital Media Technology IDMT; (<https://www.idmt.fraunhofer.de/en.html>). I recenti progressi nel campo della computer vision e del machine learning consentono di estrarre informazioni e metadati interessanti da immagini e video più velocemente di quanto potrebbe fare qualsiasi essere umano, cosa particolarmente importante con la crescente quantità di dati digitali o digitati. Fraunhofer, responsabile dello sviluppo e della configurazione degli strumenti di automatic tagging, propone l'utilizzo di diversi strumenti per eseguire in automatico la segmentazione video temporale, il rilevamento del movimento della telecamera, la misurazione della qualità video ed il rilevamento e riconoscimento degli oggetti. Tutti questi strumenti sono stati integrati nella piattaforma e orchestrati attraverso una

Fig 2
IMC pipeline



catena di procedure da parte di CINECA.

2. Annotazione e tagging manuale

Per agevolare il tagging manuale di un così vasto repertorio è stato predisposto un vocabolario controllato costituito da liste di termini a tre livelli di classificazione che coprono i topics definiti come essenziali per lo studio multidisciplinare del territorio e della storia urbana attraverso il materiale multimediale, per citarne alcuni: Architettura, Sviluppo urbano, Storia dei veicoli di trasporto, Sistema del traffico, Turismo, Eventi e Celebrazioni. Poiché il progetto I-Media-Cities adotta un approccio Linked Data per il metadata model, si sta ricercando l'eventuale versione Linked Data per ogni lista di termini specifica del vocabolario per arricchire con informazione resa disponibile da differenti domini di conoscenza. Ad esempio informazioni sullo stile architettonico possono essere annotate sia con termini dal Vocabolario del progetto sia recuperandole da un thesaurus o un'ontologia di dominio che le mette a disposizione come Linked Data. Inoltre nel sistema di annotazione è implementato il collegamento con Geonames (<http://www.geonames.org/>) per annotare le coordinate geografiche di ogni edificio importante, monumento o landmark individuato.

Il sistema di annotazione descritto fino ad ora prevede l'inserimento del tag a livello di ogni singolo segmento del video analizzato. Per i metadati sia tecnici che descrittivi che corredano ogni contenuto multimediale nel suo complesso sono stati selezionati sia lo standard di meta datazione sia i vocabolari specifici usati dallo European Film Gateway project (<http://www.europeanfilmgateway.eu/>).

L'obiettivo del modello di annotazione di I-Media-Cities è, quindi, fornire un collegamento verso i domini della conoscenza, gettando le basi per un sistema di annotazioni semantico interconnesso ad altri set Open Data disponibili via Web (come il summenzionato Geonames o Wikidata).

Naturalmente l'uso estensivo di vocabolari controllati e di servizi Linked Open Data sarà comunque affiancato dalla possibilità di aggiungere annotazioni a testo libero.

3. Interfaccia utente

Il framework back-end di I-Media-Cities fornisce funzionalità di "ricerca e browsing" che elaborano le richieste provenienti dall'interfaccia utente attraverso il front end del portale web e, conseguentemente, recuperano le informazioni relative agli elementi contenuti, che sono state caricate nel repository Neo4J (<https://neo4j.com/>), sotto forma di oggetti grafici.

I metadati raccolti possono essere presentati agli utenti attraverso diverse modalità: geolocalizzati sulle mappe oppure mostrando eventi lungo una linea temporale o presentando informazioni strutturate in un albero interattivo. Una caratteristica interessante è il "reverse storyboard" che, recuperando dal film completo l'immagine iniziale di ogni singolo shot, fornisce un riepilogo leggibile o stampabile di un filmato, incluse le istantanee di riprese, arricchite con le annotazioni di movimento di camera e la trascrizione automatica del discorso parlato, se presente.

Lo storyboard video (figura 3) include il numero della scena a cui si riferisce, l'intervallo di frame, un'immagine key frame della scena, informazioni temporali sulla scena, come start, timecode e durata e, come detto, i parametri di movimento della fotocamera relativi alla scena.

Shot Number	Frame Range	Shot Start/End	Shot Duration (sec)	Camera Motion
1	61-1220	00:00:00-00:00:50	50	no motion: 0.364
2	1221-1401	00:00:50-00:01:01	9	no motion: 0.280
3	1402-1504	00:01:01-00:01:12	4	no motion: 0.360
4	1505-1647	00:01:12-00:01:37	11	no motion: 0.390
5	1648-1819	00:01:37-00:01:49	4	no motion: 0.549
6	1900-2216	00:01:21-00:01:46	11	no motion: 0.31
7	2218-2267	00:01:35-00:01:37	3	no motion: 0.544
8	2528-2750	00:01:37-00:01:40	16	no motion: 0.258
9	2761-2976	00:01:35-00:01:41	6	no motion: 0.277
10	2978-3402	00:02:04-00:02:37	17	no motion: 0.394

Fig 3
Funzionalità di reverse storyboard che mostra le scene, i fotogrammi, ed i luoghi riferiti al materiale audiovisivo

La visualizzazione e le funzionalità messe a disposizione nel portale web di IMC saranno differenziate a seconda delle tipologie di utenza. Al momento si prevedono servizi pensati per i ricercatori e servizi dedicati al pubblico in generale. I ricercatori avranno l'opportunità di pubblicare i risultati in forma di raccolte dinamiche di contenuti A/V e immagini fisse. Lo strumento di raccolta dinamica, denominato Virtual Collection Creator, è stato progettato per creare nuovi modi di navigare e presentare i contenuti per le diverse aree di ricerca definite nel progetto. Il Virtual Collection Creator simulerà, infatti, spazi espositivi tridimensionali in cui raccogliere e presentare immagine statiche e in movimento e documenti risultato delle ricerche sulla piattaforma.

4. Conclusioni

Il progetto I-Media-Cities, sviluppato con una metodologia di tipo Agile (https://it.wikipedia.org/wiki/Metodologia_agile), proseguirà nel suo percorso iterativo con un continuo confronto fra sviluppatori della piattaforma e curatori dei contenuti, tra fasi di sviluppo e feedback, verso la realizzazione di una piattaforma in grado di usare algoritmi e crowdsourcing rispettivamente per il tagging automatico e manuale, utilizzando anche le risorse di calcolo più opportune come, per esempio, quelle basate su GPU. L'adattamento dinamico consentirà infatti anche a questo processo di procedere eventualmente in maniera ricorsiva, in grado di accogliere via via sempre nuovi algoritmi e annotazioni, per una piattaforma capace di porsi come un nuovo ecosistema digitale a disposizione di ricercatori e cittadini alla scoperta di una comune identità europea attraverso lo studio degli ambienti urbani.

Riferimenti bibliografici

Baraldi Lorenzo, Grana Costantino, Cucchiara Rita, "Recognizing and presenting the storytelling video structure with deep multimodal networks" IEEE Transactions on mul-

timedia, pp. 1 -14 , 2016

Baraldi Lorenzo, Grana Costantino, Cucchiara Rita, "Scene-driven retrieval in edited videos using aesthetic and semantic deep features" Proceedings of the 2016 ACM on International Conference on Multimedia Retrieval, New York, USA, pp. 23 -29 , 6-9 Giugno 2016, 2016

Doi: 10.1145/2911996.2912012

Capoccioni Carlo, Porcari Giovanni, il Muvir: il primo museo virtuale delle banche operanti in Italia, in Negri Clementi Gianfranco, Economia dell'arte: Proteggere, gestire e valorizzare le opere d'arte, EAGEA, Milano, 2017

Autori

Simona Caraceni s.caraceni@ Cineca.it

PhD alla Plymouth University in Aesthetics and Technology, si occupa di progetti per Musei e Beni Culturali per il Visit-Lab del Dipartimento di SuperCalcolo, Applicazioni ed Innovazione del Cineca



Michele Carpené m.carpen@ Cineca.it

Laureato presso la Facoltà di Scienze dell'Informazione all'Università degli Studi di Bologna, specializzatosi in automazione. Software Engineer con esperienza in C++, Java e PHP. Presso CINECA su progetti internazionali presso il Middleware team e nello sviluppo di applicativi interni.

Mattia D'Antonio m.dantonio@ Cineca.it

Master in Informatica presso l'Università di Sapienza (Roma, IT) e un dottorato di ricerca in Biochimica, Biologia Molecolare e Bioinformatica presso l'Università di Bari, dal 2007 lavora presso CINECA nel gruppo middleware HPC per la gestione dei dati. La sua attività principale è la costruzione e la gestione di servizi di calcolo ad alte prestazioni, database e flussi di lavoro automatizzati di analisi collaborando in numerosi progetti di ricerca nazionali ed internazionali.



Giuseppe Fiameni g.fiameni@ Cineca.it

Laureato in Scienze Informatiche presso l'Università di Bologna ed è un dottorando all'Università di Modena e Reggio Emilia. Guida il gruppo Servizi Dipartimento di SuperCalcolo, Applicazioni ed Innovazione del Cineca. Attualmente è coinvolto in diversi progetti europei, tra cui il progetto Human Brain, EUDAT2020 e PRACE, in attività di ricerca e sviluppo nel campo della gestione e dell'elaborazione dei dati.



Antonella Guidazzoli a.guidazzoli@ Cineca.it

MD in Ingegneria Elettronica; MD in Storia, "cum laude". Attività: visualizzazione scientifica, archeologia virtuale, digital cultural heritage. Responsabile del Visit Lab (Visual Information Technology Laboratory) del Dipartimento di SuperCalcolo, Applicazioni ed Innovazione del Cineca.

Silvano Imboden s.imboden@ Cineca.it

MD in Computer Science. Senior architect e sviluppatore al Cineca; supervisor tecnico della

computer grafica.

Maria Chiara Liguori m.liguori@ Cineca.it

Laureata in Scienze Politiche ed in Storia Contemporanea, con un dottorato di ricerca in Storia. Coordina progetti presso il Visit-Lab del Dipartimento di SuperCalcolo, Applicazioni ed Innovazione del Cineca.

Margherita Montanari m.montanari@ Cineca.it

Laureata in Sociologia. Lavora presso Cineca dal 1990, ha partecipato allo sviluppo di sistemi di supporto alle decisioni e applicazioni per la gestione e la rappresentazione di domini di conoscenza per gli aspetti di estrazione analisi e classificazione di concetti, predisposizione di thesauri e tassonomie. Attualmente partecipa alle attività di analisi dei requisiti , gestione metadati e test di funzionalità delle applicazioni sviluppate.



Giuseppe Trotta g.trotta@ Cineca.it

Laurea in Informatica prima di entrare nel progetto mEDRA dal 2008 occupandosi del repository metadati esistente e ai diversi servizi web costruiti su di esso. È entrato a far parte della CINECA nel 2016 nel "team metadati" del dipartimento HPC. È principalmente coinvolto nella progettazione e nello sviluppo di repository di metadati e applicazioni semantiche.

Gabriella Scipione g.scipione@ Cineca.it

Coordina da vari anni all'interno del dipartimento HPC in CINECA il team di sviluppo che si occupa principalmente di metadata management, rights management, open access e persistent identifiers, . E' responsabile da diversi anni di numerosi progetti europei, fra i quali I-Media-Cities, FORWARD (<http://project-forward.eu/>), RDI (<http://www.rdi-project.org/>)Arrow, Arrow Plus (www.arrow-net.eu), VOA3R (www.voa3r.eu), mEDRA (<https://www.medra.org/>).

