

“PH-Remix”: un progetto interdisciplinare per la valorizzazione del patrimonio audiovisivo del Festival dei Popoli-Festival Internazionale del Film Documentario di Firenze in ambiente digitale

Giovanni Grasso¹, Chiara Mannari², Davide Serramazza³

¹ Università degli Studi di Pisa, Italia - giovanni.grasso@fileli.unipi.it

² Università degli Studi di Pisa, Italia - chiara.mannari@fileli.unipi.it

³ Università degli Studi di Pisa, Italia - davide.serramazza@fileli.unipi.it

...

ABSTRACT

Il seguente contributo mira ad illustrare le finalità previste dal progetto “PH-Remix” (*Public History Remix*) dell’Università di Pisa per la valorizzazione del patrimonio filmico del Festival dei Popoli - Festival Internazionale del Film Documentario di Firenze in collaborazione con la Fondazione Mediateca Toscana (quest’ultima dal 2011 socio FIAF [International Federation of Film Archives]). Un progetto biennale ancora in fase embrionale che ha lo scopo di valorizzare, attraverso la creazione di una piattaforma web basata sull’intelligenza artificiale, l’ampio patrimonio filmico (oltre 25.000 titoli) in diversi formati (pellicole 16 e 35 mm, VHS, U-matic, DVD) raccolto e conservato nel corso dei sessant’anni di attività del Festival dei Popoli; un’eredità culturale in continua espansione grazie allo svolgimento del FdP a cui pervengono annualmente circa 1600 nuove opere.

L’archivio cinematografico del FdP ha cominciato a sedimentarsi sin dalle prime edizioni (la prima è del 1959), quando emerse la consapevolezza tra gli organizzatori di dover conservare un patrimonio culturale cinematografico distante da quelli che erano i circuiti di distribuzione piuttosto tradizionali e convenzionali del tempo. La complessità e la varietà dei titoli presenti nell’archivio riflettono l’evoluzione del cinema documentario negli ultimi sessant’anni per contenuti, stile e linguaggio (in considerazione anche degli oltre settanta paesi partecipanti), che includono le più disparate categorie del cinema documentario da quelle più strettamente antropologiche e sociali a quelle riguardanti la storia contemporanea, la storia dell’arte, la musica, la danza, il costume, l’architettura, i mezzi di comunicazione, l’attualità e l’attualità politica. Fra le varie opere presenti nell’archivio spiccano quelle dei maestri Joris Ivens, Ermanno Olmi, Vittorio De Seta, Ken Loach, Werner Herzog, Marcel Łoziński, Vincent Dieutre, Jean Rouch, Frederick Wiseman e Nagisa Oshima per citarne alcuni.

PAROLE CHIAVE

Cinema e storia; Remix culture; Intelligenza artificiale; Archivi audiovisivi; Piattaforma web; Digital Public History;

STUDIO DELLE FONTI AUDIOVISIVE NELLA DIGITAL PUBLIC HISTORY

La salvaguardia e la valorizzazione di questo ingente patrimonio nasce dalla consapevolezza di dover render note le due narrazioni di cui la storia degli audiovisivi (in questo caso dell’archivio del FdP) si compone: da un lato il prodotto filmico inteso come l’intenzionalità degli autori e la sua diffusione in qualità di testimone del mondo contemporaneo e dell’immaginario sociale; dall’altro la storia fisica delle copie che a tali prodotti hanno fornito il supporto materiale, cercando così di trasformare il passato in una risorsa e, al contempo, restituire dignità culturale a queste produzioni.

Per quanto concerne quest’ultimo aspetto risulta necessario un intervento volto ad individuare quelle copie su supporti analogici che necessitano di essere preservate sulla base del valore storico dell’opera, sulla base del rischio di deterioramento della copia fisica e sulle potenzialità del loro riutilizzo in un’ottica di *remix culture* per la creazione di nuovi prodotti audiovisivi in grado di sviluppare nuovi punti di vista, nuove storie, nuovi significati, utilizzando le fonti filmiche primarie.

Per quanto invece riguarda il valore di documento storico, il cinema non deve essere considerato soltanto come spettacolo e/o mero intrattenimento: grazie ad innovativi strumenti in grado di rendere fruibili in rete cospicui archivi audiovisivi

sedimentatisi in ambiente digitale, si deve oggi considerare il documento audiovisivo come una nuova risorsa per lo studio e la trasmissione della storia al fianco della memoria ufficiale divulgata attraverso manuali e volumi di storia tradizionali, anche di storia del cinema. Come scrive Pierre Sorlin: «I prodotti audiovisivi non sono, dunque, una documentazione supplementare, un mezzo per conoscere meglio gli eventi, i mestieri, le feste, le cerimonie, ma sono del tutto integrati nella vita, nelle normali attività, formano la trama della quotidianità, né più né meno delle case, il lavoro, le relazioni familiari. [...] pur senza esserne sempre consapevoli, noi siamo integrati in un vasto sistema di rappresentazione: osserviamo il mondo sullo schermo, ma siamo ugualmente quelli stessi che vengono filmati dalla cinepresa»¹.

Partendo da questi presupposti, lo storico che intende utilizzare il documento audiovisivo come nuova risorsa per lo studio della storia deve considerare i rapporti tra cinema e storia secondo tre prospettive principali: il cinema come agente di storia (come creatore di tendenze, immaginari e comportamenti collettivi); il cinema come strumento per raccontare la storia (strumento per raccontare il passato che il film ci racconta e rappresenta sullo schermo); ed infine, il cinema come fonte per la storia riferendosi al presente in cui il film è stato girato (all'interno di un discorso meta-cinematografico che analizza tutti quegli elementi che ci portano dentro la storia del cinema valutando il significato del film nel suo contesto produttivo, socioeconomico e politico).

Ma parallelamente a quest'approccio e alla metodologia di ricerca tradizionale, lo storico che intende servirsi del cinema per lo studio della storia, deve fare i conti con «la rivoluzione digitale [che] ha avuto un profondo impatto sul modo in cui oggi la storia viene studiata, analizzata, condivisa, insegnata oltre che sul modo in cui le fonti che documentano i fenomeni del passato vengono pubblicate, conservate e addirittura prodotte». Lo storico deve necessariamente riflettere sul proprio ruolo sociale nel presente, sull'uso pubblico della storia e del cinema da parte dei soggetti sociali e superare una certa diffidenza «verso il computer e la rete, lo scarso valore attribuito ai prodotti digitali, [...] l'ignoranza diffusa relativa ai principali strumenti del mondo digitale, la loro rapida e continua evoluzione» e abbandonare l'opinione «che l'informatica sia appunto solo uno 'strumento', potenzialmente utile all'umanista solo nella misura in cui lo aiuta a risolvere alcune esigenze di ricerca e non un modo di analizzare, organizzare e visualizzare la conoscenza»².

Il progetto nasce proprio dall'esigenza di creare un nuovo strumento per raccontare e rielaborare la storia degli audiovisivi che compongono l'archivio del FdP sotto i due punti di vista poc'anzi presentati e fare in modo che diverse tipologie di pubblico possano usufruire in rete di un patrimonio archivistico tra i più importanti in ambito audiovisivo; ciò è possibile unendo tre ambiti disciplinari differenti (la storiografia cinematografica, l'informatica e l'informatica umanistica), condividendo i metodi e le strategie derivanti dai diversi settori e giungendo ad una consapevolezza delle metodologie delle tre discipline: multidisciplinarietà, o meglio interdisciplinarietà, che è alla base della *digital public history*.

LA PIATTAFORMA “PH-REMIX”

La piattaforma “PH-Remix” che nasce e si sviluppa all'interno dell'omonimo progetto è lo strumento previsto per consentire l'accesso alle fonti e il remix dei video. Il patrimonio filmico del FdP costituisce il nucleo di partenza che andrà ad alimentare la piattaforma ed è l'oggetto del processo di estrazione automatica presentato in seguito. L'archivio, al cui interno troviamo film documentari provenienti da ogni parte del mondo che affrontano temi disparati con una eterogeneità di stili ed approcci, valorizza autori lontani dal centro del mercato cinematografico e si presta bene allo studio e al riutilizzo dei materiali ai fini della produzione di nuovi contenuti di tema storico.

La piattaforma, tuttora in corso di progettazione, sarà un'applicazione web composta da due componenti essenziali: la prima componente fornirà l'accesso ad un indice molto ampio da cui sarà possibile, attraverso percorsi di ricerca e navigazione, visionare ed infine selezionare i contenuti accessibili sotto forma di clip. Le clip consistono in frammenti video di durata variabile contenenti scene o momenti significativi estratti, oltre che attraverso una schedatura tradizionale basata su temi e parole chiave, anche tramite metodi innovativi di estrazione automatica dell'informazione ottenuti mediante l'uso di tecniche di intelligenza artificiale; la seconda componente si occuperà del montaggio del nuovo video attraverso l'unione di più clip, l'aggiunta di immagini, testi, tracce audio e l'esportazione finale.

Una piattaforma di questo tipo si inserisce nel contesto della *remix culture*, una pratica di produzione artistica basata sul riutilizzo di contenuti provenienti da svariate fonti che ha preso piede negli anni Sessanta nel panorama dei remix musicali per poi diffondersi anche in campo audiovisivo. Con l'invenzione del web la *remix culture* ha conquistato terreno e sono

¹ Sorlin, P., *L'immagine e l'evento. L'uso storico delle fonti audiovisive*, Paravia, 1999, p. 6.

² Salvatori, E., *Digital (Public) History: la nuova strada di una antica disciplina*, in Sini G. (a cura di), *Scienze umane, dalla produzione di nuova conoscenza alla disseminazione e ritorno*, “RiMe — Rivista dell'Istituto di Storia dell'Europa Mediterranea”, n. 1, 2017, pp. 57-59.

stati aperti numerosi siti web e spazi virtuali dove condividere progetti di remix³. Nell'era di Instagram Reels⁴, delle coreografie su Tik Tok⁵ e dei film interattivi⁶, in una realtà in cui le tradizionali modalità di accesso ai contenuti sono stravolte e gli utenti sono diventati allo stesso tempo fruitori e creatori, la *remix culture* è penetrata al punto da far perdere la percezione di sé: *everything is a Remix*⁷.

La natura interdisciplinare del progetto “PH-Remix” permette di applicare buone pratiche ed elementi innovativi provenienti dalle discipline coinvolte facendo emergere i seguenti punti di forza:

- Sperimentare la pratica del remix nell'ambito della *digital public history* in quanto disciplina che ha come obiettivo primario quello di avvicinare storia e pubblico attraverso lo sviluppo di strumenti al passo con l'evoluzione tecnologica e di nuove modalità di interazione;
- Rendere accessibile agli studiosi fonti storiche, la cui fruizione risultava più complicata, grazie all'accesso a un ingente patrimonio filmico digitalizzato, schedato e indicizzato in maniera sistematica;
- Aprire ad una serie di scenari interessanti per la didattica e la produzione artistica;
- Servirsi del contributo dell'intelligenza artificiale, delle tecniche di estrazione automatica e in particolar modo del *machine learning* per la valorizzazione del patrimonio storico e culturale grazie alla possibilità di processare, in tempi brevi, quantità di dati impossibili da trattare con schedature manuali limitate nell'azione da vincoli temporali.

Le problematiche emerse fin dalle fasi iniziali del progetto riguardano soprattutto due aspetti:

- Il primo consiste nella tutela giuridica dei materiali d'archivio e lo sviluppo della piattaforma nel rispetto del diritto d'autore. In questa direzione si sta tentando, in accordo con i vari partner di progetto, di impiegare un sistema di licenze di utilizzo che si basi sul modello delle *Creative Commons*⁸;
- Il secondo riguarda alcune difficoltà nella definizione delle specifiche della piattaforma dipendenti sia dalle tempistiche interne al progetto “PH-Remix”, in particolare in merito ai tempi necessari all'analisi dell'archivio FdP e alla sperimentazione delle tecniche di estrazione dell'informazione, che dalle tempistiche del progetto parallelo di Fondazione Mediateca Toscana che prevede l'acquisizione dell'archivio FdP all'interno della mediateca e la digitalizzazione dell'intero patrimonio archivistico (i cui materiali sono al momento conservati in copia unica su supporto analogico).

Per far fronte a queste difficoltà, emerge quindi la necessità di creare un sistema flessibile in cui le modifiche di progetto siano gestite in maniera ottimale anche sul lungo termine. La soluzione per soddisfare questa esigenza è quella di realizzare un'architettura software basata su microservizi in cui tutte le componenti siano separate e implementabili anche in momenti successivi. Il sistema verrà progettato affinché risulti possibile in qualsiasi momento aggiungere agevolmente nuove funzionalità come ad esempio nuovi criteri di indicizzazione e di ricerca, oppure sottomettere e analizzare altri film in maniera piuttosto semplice anche tramite un'interfaccia web utilizzabile da utenti non esperti. L'architettura prevista permette inoltre di avere maggiore flessibilità per quanto riguarda lo strumento di remix: l'indice e tutti i moduli che verranno implementati potranno essere riutilizzati all'interno di nuove interfacce, come ad esempio nel caso in cui si decida di affiancare all'applicazione web in corso di sviluppo una o più app mobile.

IL CONTRIBUTO DELL'INTELLIGENZA ARTIFICIALE

Per tale tipo di patrimonio composto da materiale audiovisivo, la piattaforma sin qui descritta può essere realizzata mediante l'uso di tecniche d'intelligenza artificiale da affiancare a un processo più 'classico' di catalogazione manuale.

In letteratura scientifica sono presenti diversi modelli pre-allenati per etichettare parti di video: questi modelli, da qui in poi chiamati anche *processor*, possono essere assoggettati agli scopi del progetto per suddividere i documentari in varie

³ Cfr. *Rip! A Remix manifesto*: <https://www.brettgaylor.com/#/rip-a-remix-manifesto/>

⁴ <https://about.instagram.com/blog/announcements/introducing-instagram-reels-announcement>

⁵ <https://www.nytimes.com/2019/03/10/style/what-is-tik-tok.html>

⁶ <https://medium.com/swlh/films-that-watch-you-while-you-watch-them-a-brief-history-of-interactive-movies-820fb5e2ee25>

⁷ Cfr. Ferguson K., *Everything is a Remix*: <https://www.everythingsaremix.info/watch-the-series> (per tutti: ultima consultazione sett. 2020).

⁸ <https://creativecommons.org/>

clip, ovvero in frammenti contigui e omogenei all'interno del film. Le clip così estratte saranno inoltre **indicizzate**, assegnando ad ognuna di esse una **label** che andrà ad indicare l'informazione ricavata sulla porzione di video dai modelli adoperati.

Per entrare nel dettaglio dell'intero processo composto da estrazione, indicizzazione ed etichettamento, poniamo un esempio: possibili *task* svolti dai *processor* consistono nell'identificazione di oggetti nelle clip come in *Looking fast and slow*⁹, di azioni compiute come nell'algorithmo di Kalfaoglu¹⁰ o ancora nell'estrazione di testo dal video (in questo caso i sottotitoli) come nel lavoro di Wang¹¹. I *processor* vengono utilizzati per suddividere un singolo film in più clip basandosi sui *task* precedentemente menzionati. Inoltre, è da sottolineare come da un singolo film vengano estratti **più insiemi di clip**, uno per ciascun *processor* utilizzato; fra clip estratte dallo stesso documentario possono potenzialmente verificarsi delle sovrapposizioni. Infatti, una clip in cui appare l'oggetto A, si può sovrapporre, anche solo parzialmente, ad un'altra clip in cui è compiuta l'azione B, ma anche ad un'altra clip in cui compare un altro oggetto C.

Una volta isolate dal resto del documentario, le clip riceveranno una *label* corrispondente all'informazione estratta dal modello utilizzato: le *label* assegnate coincideranno con le parole chiave utilizzate dagli utenti della piattaforma di remix per **ricercare** una particolare clip (ad esempio tutte quelle contenenti un particolare tipo di oggetto, di azione, ecc.). Tale processo andrà a costruire l'indice a cui si è accennato nella sezione precedente, che sarà composto dalle varie clip estratte dai film a nostra disposizione e dalle *label* ad esse associate. Gli esempi riportati precedentemente di *processor* non sono da ritenersi esaustivi: il numero, le categorie alle quali apparterranno e i modelli specifici che verranno adoperati, saranno decisi durante lo svolgimento del progetto.

Ai modelli pre-allenati saranno affiancati altri tipi di modelli sviluppati *ex novo*: verrà dunque interamente progettata l'architettura e saranno successivamente allenati usando i documentari come dati di allenamento. Quest'ultimo tipo di modelli risulta ancora più interessante dal punto di vista della ricerca scientifica.

Un esempio di tali funzionalità riguarda la possibilità da parte dell'utente finale, una volta scelte due clip A e C, di poter ottenere dal sistema un suggerimento di una terza clip B come intermezzo adeguato: in tal modo si potrebbero scoprire relazioni non banali tra clip provenienti da due film con tematiche non apparentemente collegate tra loro.

Tecnicamente ciò potrebbe essere realizzato tramite tecniche di estrazione di *features* a partire dalle clip e tramite l'utilizzo di uno spazio comune di *embedding*: le varie clip precedentemente estratte saranno dunque mappate in uno **spazio semantico comune** nel quale sarà possibile calcolare la distanza fra i vari elementi. A tal proposito il nostro intento è quello di applicare al video l'utilizzo degli *embedding* neurali per modellare ed estrarre nuove relazioni (già investigate nel lavoro di Xu)¹² e sfruttarli per quantificare la similarità tra le varie clip e dunque scoprire nuove relazioni fra esse.

Un altro esempio di funzionalità sviluppata *ex novo* sarà il supporto fornito agli esperti del settore nella catalogazione dei film: il modello previsto, analizzando aspetti quali il contenuto delle immagini, delle musiche di sottofondo e dei dialoghi di una determinata sequenza, predirà le tematiche trattate. La problematica maggiore risulta essere l'identificazione dell'insieme delle tematiche che tale modello può predire: se i documentari hanno per la maggior parte un'indicazione delle tematiche trattate verranno utilizzate tali categorie per comporre il nostro insieme, altrimenti il modello prevederà l'uso di tecniche *unsupervised* che non richiedono la presenza di dati etichettati.

Quanto descritto in questa sezione, ovvero sia i *processor* che i nuovi modelli che ci si propone di sviluppare, verranno tutti implementati come eseguibili separati nell'ottica di costruire un'architettura a microservizi come precedentemente accennato. I vantaggi di tale scelta sono principalmente due: ogni funzionalità è indipendente dalle altre, ovvero l'aggiunta/rimozione/modifica di una singola funzionalità sarà molto più agevole rispetto a un corrispettivo monolitico e in secondo luogo viene favorita la scalabilità del sistema in maniera tale da poter sfruttare tutte le risorse computazionali in nostro possesso.

I risultati di tale sezione del progetto, oltre a essere indispensabili per la piattaforma "PH-Remix" alla quale verranno fornite le clip ricercabili dagli utenti, sono di per sé utili anche ad altri tipi di studi e analisi da svolgere su di essi. Un esempio su tutti riguarda la trascrizione dei dialoghi presenti nel testo: da esse è possibile ottenere dei corpora etichettati con le varie informazioni ricavabili dai documentari, sia in maniera automatica con gli strumenti descritti in precedenza, che manualmente con l'ausilio di un esperto. Questi corpora costituirebbero una risorsa preziosa per alcuni studiosi, come

⁹ Liu, M., *Looking fast and slow: Memory-guided mobile video object detection*, preprint arXiv: 1903.10172, 2019.

¹⁰ Kalfaoglu, E., Kalkan, S., Alatan, A., *Late Temporal Modeling in 3D CNN Architectures with BERT for Action Recognition*, preprint arXiv: 2008.01232, 2020.

¹¹ Wang, T., *End-to-end text recognition with convolutional neural networks*, Proceedings of the 21st international conference on pattern recognition (ICPR2012), IEEE, 2012.

¹² Xu, P., Denilson, B., *Investigations on knowledge base embedding for relation prediction and extraction*, arXiv preprint arXiv: 1802.02114, 2018.

ad esempio storici o linguisti, che potranno ulteriormente analizzare tale corpora autonomamente in maniera del tutto indipendente da tale progetto, utilizzando altre metodologie più consone alle loro discipline.

BIBLIOGRAFIA

- [1] Bell, Desmond. «Documentary film and the poetics of history», *Journal of Media Practice* (2011).
- [2] De Luna, Giovanni. «Le nuove frontiere della storia. Il cinema come documento storico», *Le fonti - audiovisive per la storia e la didattica*, Annali AAMOD n. 16 (2013).
- [3] Gallagher, Owen. «Reclaiming critical remix: the role of sampling in transformative works», Routledge (2018).
- [4] Kalfaoglu, Esat, Kalkan, Sinan, Alatan, Aydin. «Late Temporal Modeling in 3D CNN Architectures with BERT for Action Recognition», arXiv preprint arXiv: 2008.01232 (2020).
- [5] Lagny, Michèle. «Il cinema come fonte di storia», in Brunetta Gian Piero (a cura di), *Storia del cinema mondiale*, Vol. 5 Teorie, strumenti, memorie, Einaudi (2001).
- [6] Liu, Mason, et al. «Looking fast and slow: Memory-guided mobile video object detection», arXiv preprint arXiv: 1903.10172 (2019).
- [7] Navas, Eduardo. «Remix theory. The aesthetics of sampling», Springer (2012).
- [8] Pang, Natalie, Liew, Kai Khiun, Chan, Brenda, «Participatory archives in a world of ubiquitous media», *Archives and Manuscripts* (2014).
- [9] Rosenstone, A. Robert. «Revisioning History. Film and the Construction of a New Past», Princeton University Press (1995).
- [10] Salvatori, Enrica. «Digital (Public) History: la nuova strada di una antica disciplina», in Sini Giovanni (a cura di), *Scienze umane, dalla produzione di nuova conoscenza alla disseminazione e ritorno*, “RiMe — Rivista dell’Istituto di Storia dell’Europa Mediterranea” n. 1 (2017).
- [11] Sorlin, Pierre. «L’immagine e l’evento. L’uso storico delle fonti audiovisive», Paravia (1999).
- [12] Sorlin, Pierre. «Una premessa. Scuola, storia, mezzi audiovisivi nell’era digitale». *Le fonti - audiovisive per la storia e la didattica*, Annali AAMOD n. 16 (2013).
- [13] Taviani, Ermanno, «Il cinema e gli storici». *Le fonti - audiovisive per la storia e la didattica*, Annali AAMOD n. 16 (2013).
- [14] Wang, Tao. «End-to-end text recognition with convolutional neural networks». *Proceedings of the 21st international conference on pattern recognition (ICPR2012)*, IEEE (2012).
- [15] Xu, Peng, Denilson, Barbosa. «Investigations on knowledge base embedding for relation prediction and extraction». arXiv preprint arXiv: 1802.02114 (2018).